[redacted]

GenCore version 5.1.3

...otein - protein search, using sw model

Run on: January 14, 2003, 18:39:18 ; Search time 14 Seconds
(without alignments)
39.931 Million cell updates/sec

Title: US-09-828-645-1
Perfect score: 119
Sequence: 1 DICNTMHYTNWTHIYICEE 19

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	119	100.0	20	2	US-08-934-915-28	Sequence 28, Appl
2	119	100.0	365	2	US-08-472-666-4	Sequence 4, Appli
3	119	100.0	365	4	US-09-362-012A-2	Sequence 2, Appli
4	119	100.0	365	4	US-09-362-012A-5	Sequence 5, Appli
5	119	100.0	365	5	PCT-US96-07615-4	Sequence 4, Appli
6	114	95.8	20	2	US-08-934-915-178	Sequence 178, App
7	109	91.6	20	2	US-08-934-915-31	Sequence 31, Appl
8	104	87.4	20	2	US-08-934-915-136	Sequence 136, App
9	96	80.7	20	2	US-08-934-915-177	Sequence 177, App
10	79	66.4	20	2	US-08-934-915-69	Sequence 69, Appl
11	57	47.9	20	2	US-08-934-915-74	Sequence 74, Appl
12	52	43.7	20	2	US-08-934-915-13	Sequence 13, Appl
13	47	39.5	20	2	US-08-934-915-7	Sequence 7, Appli
14	47	39.5	1587	4	US-09-000-094-46	Sequence 46, Appl
15	44	37.0	19	4	dimeric-glycoprotein-rec	Sequence 2, Appli
16	43	36.1	257	2	US-08-685-992-6	Sequence 6, Appli
17	43	36.1	257	2	US-09-144-925-6	Sequence 6, Appli
18	42	35.3	20	2	US-08-934-915-32	Sequence 32, Appl
19	42	35.3	154	4	US-09-134-001C-2992	Sequence 2992, Ap
20	42	35.3	251	2	US-08-685-992-15	Sequence 15, Appl
21	42	35.3	251	2	US-09-144-925-15	Sequence 15, Appl
22	42	35.3	489	4	US-09-134-001C-2920	Sequence 2920, Ap
23	42	35.3	500	6	5171684-2	Patent No. 5171684
24	41	34.5	236	3	US-08-705-771-19	Sequence 19, Appl
25	41	34.5	265	2	US-08-807-044-1	Sequence 1, Appli
26	41	34.5	600	2	US-08-370-156-4	Sequence 4, Appli
27	41	34.5	600	3	US-08-814-095-4	Sequence 4, Appli
28	41	34.5	600	4	US-08-975-084-1	Sequence 1, Appli
29	41	34.5	614	1	US-07-732-962A-2	Sequence 2, Appli
30	41	34.5	614	2	US-08-370-156-2	Sequence 19, Appl
31	41	34.5	614	3	US-08-446-100-19	Sequence 20, Appl
32	41	34.5	614	3	US-08-446-100-20	Sequence 21, Appl
33	41	34.5	614	3	US-08-446-100-21	Sequence 22, Appl
34	41	34.5	614	3	US-08-446-100-22	Sequence 23, Appl
35	41	34.5	614	3	US-08-446-100-23	Sequence 25, Appl
36	41	34.5	614	3	US-08-446-100-25	Sequence 2, Appli
37	41	34.5	614	3	US-08-814-095-2	Sequence 2, Appli
38	41	34.5	614	5	PCT-US92-06106-2	Sequence 6, Appli
39	41	34.5	617	2	US-08-370-156-6	Sequence 6, Appli
40	41	34.5	617	3	US-08-814-095-6	Sequence 80, Appl
41	40	33.6	238	4	US-09-257-179-80	Sequence 60, Appl
42	40	33.6	293	1	US-08-118-270-60	Sequence 60, Appl
43	40	33.6	293	5	PCT-US93-08528-60	Sequence 62, Appl
44	40	33.6	313	3	US-08-926-842B-62	Sequence 6, Appli
45	40	33.6	333	3	US-08-988-876-6	

ALIGNMENTS

RESULT 1
US-08-934-915-28
; Sequence 28, Application US/08934915
; Patent No. 5932412
; GENERAL INFORMATION:
; APPLICANT: DILLNER, JOAKIM
; APPLICANT: DILLNER, LENA
; APPLICANT: CHENG, HWEE-MING
; TITLE OF INVENTION: SYNTHETIC PEPTIDES OF HUMAN
; TITLE OF INVENTION: PAPILLOMAVIRUS 1, 5, 6, 8,
; TITLE OF INVENTION: 11, 16, 18, 31, 33 AND 56,
; TITLE OF INVENTION: USEFUL IN IMMUNOASSAY FOR
; TITLE OF INVENTION: DIAGNOSTIC PURPOSES
; NUMBER OF SEQUENCES: 193
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MASON & ASSOCIATES, P.A.
; STREET: 17757 U.S. HWY. 19 NORTH, SUITE 500
; CITY: CLEARWATER
; STATE: FLORIDA
; COUNTRY: U.S.A.
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: Windows 3.0
; SOFTWARE: Microsoft Word 6.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/934,915
; FILING DATE: 22-SEP-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/949,836
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: LOUISE A. Foutch
; REGISTRATION NUMBER: 37,133
; REFERENCE/DOCKET NUMBER: 1946.6
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 813-538-3800
; TELEFAX: 813-538-3820
; TELEX:
; INFORMATION FOR SEQ ID NO: 28:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 20 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-934-915-28

Query Match 100.0%; Score 119; DB 2; Length 20;
Best Local Similarity 100.0%; Pred. No. 3e-11;

```
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 DICNTMHYTNWTHIYICEE 19
           ||||||||||||||||||||
Db       2 DICNTMHYTNWTHIYICEE 20


RESULT 2
US-08-472-666-4
; Sequence 4, Application US/08472666
; Patent No. 5821048
;  GENERAL INFORMATION:
;    APPLICANT:  Howley, Peter M.
;    APPLICANT:  Benson, John D.
;    APPLICANT:  Yasugi, Toshiharu
;    APPLICANT:  Sakai, Hiroyuki
;    TITLE OF INVENTION:  METHOD AND KIT FOR DIAGNOSING
;    TITLE OF INVENTION:  PAPILLOMAVIRUS-INFECTED CELLS
;    NUMBER OF SEQUENCES:  4
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Ann-Louise Kerner, Ph.D.
;      ADDRESSEE:  Lappin & Kusmer
;      STREET:  200 State Street
;      CITY:  Boston
;      STATE:  MA
;      COUNTRY:  USA
;      ZIP:  01209
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/472,666
;      FILING DATE:
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  McDaniels, Patricia A.
;      REGISTRATION NUMBER:  33,194
;      REFERENCE/DOCKET NUMBER:  HAZ-010
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617-330-1300
;      TELEFAX:  617-330-1311
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  365 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;    ORIGINAL SOURCE:
;      ORGANISM:  Human papillomavirus-16 E2
;      STRAIN:  HPV-16 E2
US-08-472-666-4

  Query Match             100.0%;  Score 119;  DB 2;  Length 365;
  Best Local Similarity   100.0%;  Pred. No. 7e-10;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 DICNTMHYTNWTHIYICEE 19
           ||||||||||||||||||||
Db     124 DICNTMHYTNWTHIYICEE 142


RESULT 3
US-09-362-012A-2
; Sequence 2, Application US/09362012A
; Patent No. 6432926
;  GENERAL INFORMATION:
;       APPLICANT: Howley, Peter M.
;                  Dowhanick-Morrissette, Jennifer J.
;                  Benson, John D.
;                  Sakai, Hiroyuki
;       TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING
;                           PAPILLOMAVIRUS-INFECTED CELLS
;       NUMBER OF SEQUENCES: 5
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: LAHIVE & COCKFIELD, LLP
;            STREET: 28 State Street
;            CITY: Boston
;            STATE: Massachusetts
;            COUNTRY: USA
;            ZIP: 02109
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Floppy disk
;            COMPUTER: IBM PC compatible
;            OPERATING SYSTEM: PC-DOS/MS-DOS
;            SOFTWARE: PatentIn Release #1.0, Version #1.25
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/09/362,012A
;            FILING DATE: 27-Jul-1999
;            CLASSIFICATION: <Unknown>
;       PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: 08/677,206
;            FILING DATE: 09 JUL 1996
;       ATTORNEY/AGENT INFORMATION:
;            NAME: Jane E. Remillard
;            REGISTRATION NUMBER: 38,872
;            REFERENCE/DOCKET NUMBER: HMI-017CP
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: (617)227-7400
;            TELEFAX: (617)742-4214
;  INFORMATION FOR SEQ ID NO: 2:
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 365 amino acids
;            TYPE: amino acid
;            TOPOLOGY: linear
;       MOLECULE TYPE: protein
;       SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-362-012A-2

  Query Match             100.0%;  Score 119;  DB 4;  Length 365;
  Best Local Similarity   100.0%;  Pred. No. 7e-10;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 DICNTMHYTNWTHIYICEE 19
           ||||||||||||||||||||
Db     124 DICNTMHYTNWTHIYICEE 142


RESULT 4
US-09-362-012A-5
; Sequence 5, Application US/09362012A
; Patent No. 6432926
;  GENERAL INFORMATION:
;       APPLICANT: Howley, Peter M.
;                  Dowhanick-Morrissette, Jennifer J.
;                  Benson, John D.
;                  Sakai, Hiroyuki
;       TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING
;                           PAPILLOMAVIRUS-INFECTED CELLS
;       NUMBER OF SEQUENCES: 5
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: LAHIVE & COCKFIELD, LLP
;            STREET: 28 State Street
;            CITY: Boston
;            STATE: Massachusetts
;            COUNTRY: USA
;            ZIP: 02109
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Floppy disk
;            COMPUTER: IBM PC compatible
;            OPERATING SYSTEM: PC-DOS/MS-DOS
;            SOFTWARE: PatentIn Release #1.0, Version #1.25
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER  US/09/362,012A
;            FILING DATE  27-Jul-1999
```

; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/677,206
; FILING DATE: 09 JUL 1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Jane E. Remillard
; REGISTRATION NUMBER: 38,872
; REFERENCE/DOCKET NUMBER: HMI-017CP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)742-4214
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 365 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 5:
US-09-362-012A-5

Query Match 100.0%; Score 119; DB 4; Length 365;
Best Local Similarity 100.0%; Pred. No. 7e-10;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DICNTMHYTNWTHIYICEE 19
||||||||||||||||||||
Db 124 DICNTMHYTNWTHIYICEE 142

RESULT 5
PCT-US96-07615-4
; Sequence 4, Application PC/TUS9607615
; GENERAL INFORMATION:
; APPLICANT: PRESIDENT AND FELLOWS OF HARVARD COLLEGE
; TITLE OF INVENTION: METHODS, KITS, AND COMPOSITIONS FOR DIAGNOSING
; TITLE OF INVENTION: PAPILLOMAVIRUS INFECTION
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lappin & Kusmer
; STREET: 200 State Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 01209
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US96/07615
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Kerner, Ann-Louise
; REGISTRATION NUMBER: 33,523
; REFERENCE/DOCKET NUMBER: HAZ-010PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-330-1300
; TELEFAX: 617-330-1311
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 365 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; ORIGINAL SOURCE:
; ORGANISM: Human papillomavirus-16 E2
; STRAIN: HPV-16 E2
PCT-US96-07615-4

Query Match 100.0%; Score 119; DB 5; Length 365;
Best Local Similarity 100.0%; Pred. No. 7e-10;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DICNTMHYTNWTHIYICEE 19
||||||||||||||||||||
Db 124 DICNTMHYTNWTHIYICEE 142

RESULT 6
US-08-934-915-178
; Sequence 178, Application US/08934915
; Patent No. 5932412
; GENERAL INFORMATION:
; APPLICANT: DILLNER, JOAKIM
; APPLICANT: DILLNER, LENA
; APPLICANT: CHENG, HWEE-MING
; TITLE OF INVENTION: SYNTHETIC PEPTIDES OF HUMAN
; TITLE OF INVENTION: PAPILLOMAVIRUS 1, 5, 6, 8,
; TITLE OF INVENTION: 11, 16, 18, 31, 33 AND 56,
; TITLE OF INVENTION: USEFUL IN IMMUNOASSAY FOR
; TITLE OF INVENTION: DIAGNOSTIC PURPOSES
; NUMBER OF SEQUENCES: 193
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MASON & ASSOCIATES, P.A.
; STREET: 17757 U.S. HWY. 19 NORTH, SUITE 500
; CITY: CLEARWATER
; STATE: FLORIDA
; COUNTRY: U.S.A.
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: Windows 3.0
; SOFTWARE: Microsoft Word 6.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/934,915
; FILING DATE: 22-SEP-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/949,836
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: LOUISE A. Foutch
; REGISTRATION NUMBER: 37,133
; REFERENCE/DOCKET NUMBER: 1946.6
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 813-538-3800
; TELEFAX: 813-538-3820
; TELEX:
; INFORMATION FOR SEQ ID NO: 178:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 20 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-934-915-178

Query Match 95.8%; Score 114; DB 2; Length 20;
Best Local Similarity 94.7%; Pred. No. 1.6e-10;
Matches 18; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DICNTMHYTNWTHIYICEE 19
||||||||||||||:||||
Db 2 DICNTMHYTNWTHIWICEE 20

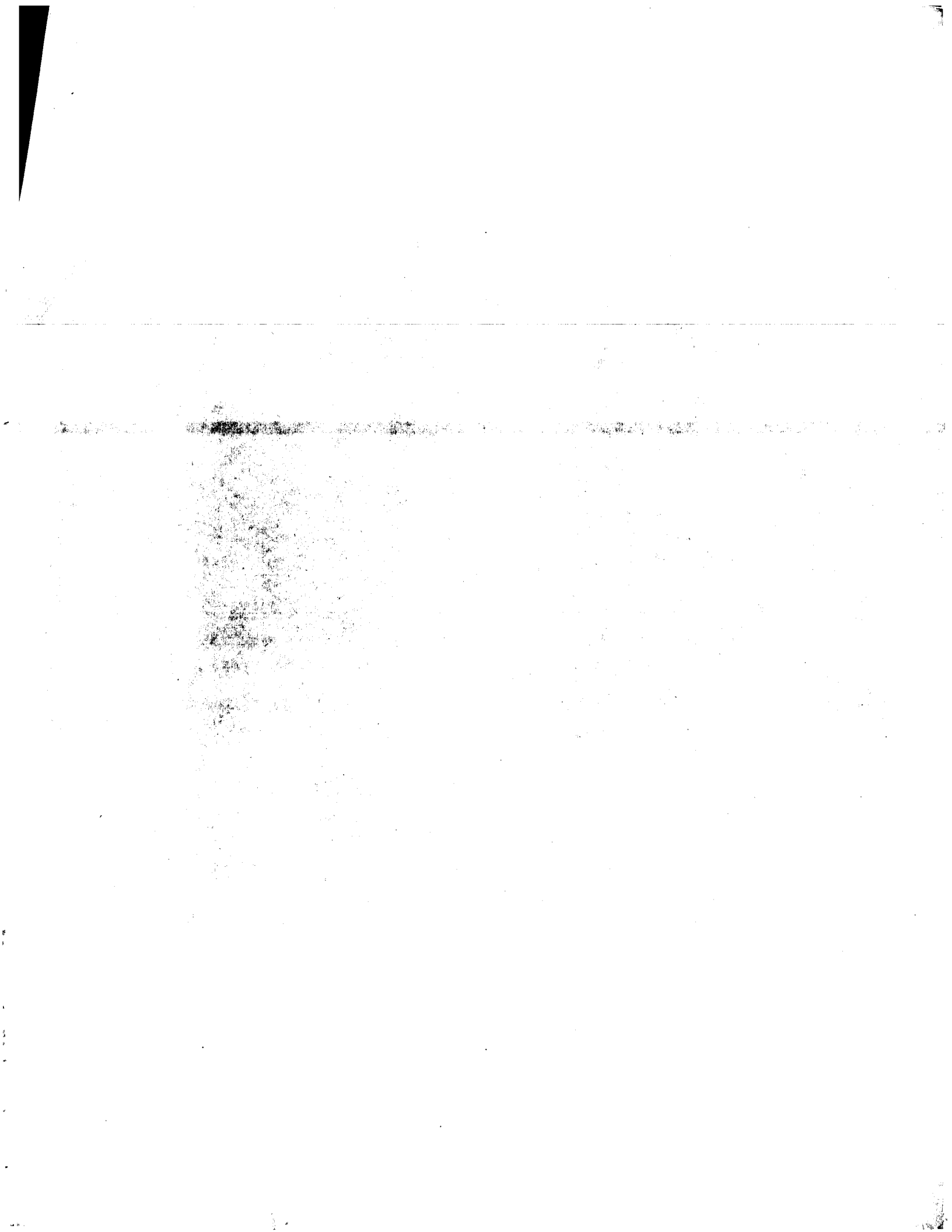
RESULT 7
US-08-934-915-31
; Sequence 31, Application US/08934915
; Patent No. 5932412
; GENERAL INFORMATION:
; APPLICANT: DILLNER, JOAKIM
; APPLICANT: DILLNER, LENA
; APPLICANT: CHENG, HWEE-MING
; TITLE OF INVENTION: SYNTHETIC PEPTIDES OF HUMAN

```
;    TITLE OF INVENTION:  PAPILLOMAVIRUS 1, 5, 6, 8,
;    TITLE OF INVENTION:  11, 16, 18, 31, 33 AND 56,
;    TITLE OF INVENTION:  USEFUL IN IMMUNOASSAY FOR
;    TITLE OF INVENTION:  DIAGNOSTIC PURPOSES
;    NUMBER OF SEQUENCES:  193
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  MASON & ASSOCIATES, P.A.
;      STREET:  17757 U.S. HWY. 19 NORTH, SUITE 500
;      CITY:  CLEARWATER
;      STATE:  FLORIDA
;      COUNTRY:  U.S.A.
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  Windows 3.0
;      SOFTWARE:  Microsoft Word 6.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/934,915
;      FILING DATE:  22-SEP-1997
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/949,836
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  LOUISE A. Foutch
;      REGISTRATION NUMBER:  37,133
;      REFERENCE/DOCKET NUMBER:  1946.6
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  813-538-3800
;      TELEFAX:  813-538-3820
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  31:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  20 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-934-915-31

  Query Match             91.6%;  Score 109;  DB 2;  Length 20;
  Best Local Similarity   100.0%;  Pred. No. 8.5e-10;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 DICNTMHYTNWTHIYIC 17
           |||||||||||||||||
Db       4 DICNTMHYTNWTHIYIC 20


RESULT 8
US-08-934-915-136
; Sequence 136, Application US/08934915
; Patent No. 5932412
;  GENERAL INFORMATION:
;    APPLICANT:  DILLNER, JOAKIM
;    APPLICANT:  DILLNER, LENA
;    APPLICANT:  CHENG, HWEE-MING
;    TITLE OF INVENTION:  SYNTHETIC PEPTIDES OF HUMAN
;    TITLE OF INVENTION:  PAPILLOMAVIRUS 1, 5, 6, 8,
;    TITLE OF INVENTION:  11, 16, 18, 31, 33 AND 56,
;    TITLE OF INVENTION:  USEFUL IN IMMUNOASSAY FOR
;    TITLE OF INVENTION:  DIAGNOSTIC PURPOSES
;    NUMBER OF SEQUENCES:  193
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  MASON & ASSOCIATES, P.A.
;      STREET:  17757 U.S. HWY. 19 NORTH, SUITE 500
;      CITY:  CLEARWATER
;      STATE:  FLORIDA
;      COUNTRY:  U.S.A.
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  Windows 3.0
;      SOFTWARE:  Microsoft Word 6.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/934,915
;      FILING DATE:  22-SEP-1997
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/949,836
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  LOUISE A. Foutch
;      REGISTRATION NUMBER:  37,133
;      REFERENCE/DOCKET NUMBER:  1946.6
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  813-538-3800
;      TELEFAX:  813-538-3820
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  136:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  20 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-934-915-136

  Query Match             87.4%;  Score 104;  DB 2;  Length 20;
  Best Local Similarity   94.1%;  Pred. No. 4.5e-09;
  Matches   16;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 DICNTMHYTNWTHIYIC 17
           :|||||||||||||||||
Db       4 NICNTMHYTNWTHIYIC 20


RESULT 9
US-08-934-915-177
; Sequence 177, Application US/08934915
; Patent No. 5932412
;  GENERAL INFORMATION:
;    APPLICANT:  DILLNER, JOAKIM
;    APPLICANT:  DILLNER, LENA
;    APPLICANT:  CHENG, HWEE-MING
;    TITLE OF INVENTION:  SYNTHETIC PEPTIDES OF HUMAN
;    TITLE OF INVENTION:  PAPILLOMAVIRUS 1, 5, 6, 8,
;    TITLE OF INVENTION:  11, 16, 18, 31, 33 AND 56,
;    TITLE OF INVENTION:  USEFUL IN IMMUNOASSAY FOR
;    TITLE OF INVENTION:  DIAGNOSTIC PURPOSES
;    NUMBER OF SEQUENCES:  193
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  MASON & ASSOCIATES, P.A.
;      STREET:  17757 U.S. HWY. 19 NORTH, SUITE 500
;      CITY:  CLEARWATER
;      STATE:  FLORIDA
;      COUNTRY:  U.S.A.
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  Windows 3.0
;      SOFTWARE:  Microsoft Word 6.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/934,915
;      FILING DATE:  22-SEP-1997
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/949,836
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  LOUISE A. Foutch
;      REGISTRATION NUMBER:  37,133
;      REFERENCE/DOCKET NUMBER:  1946.6
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  813-538-3800
;      TELEFAX:  813-538-3820
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  177:
```

```
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  20 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-934-915-177

  Query Match             80.7%;  Score 96;  DB 2;  Length 20;
  Best Local Similarity   100.0%;  Pred. No. 6.5e-08;
  Matches   15;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 DICNTMHYTNWTHIY 15
           ||||||||||||||||
Db       6 DICNTMHYTNWTHIY 20


RESULT 10
US-08-934-915-69
; Sequence 69, Application US/08934915
; Patent No. 5932412
;  GENERAL INFORMATION:
;    APPLICANT:  DILLNER, JOAKIM
;    APPLICANT:  DILLNER, LENA
;    APPLICANT:  CHENG, HWEE-MING
;    TITLE OF INVENTION:  SYNTHETIC PEPTIDES OF HUMAN
;    TITLE OF INVENTION:  PAPILLOMAVIRUS 1, 5, 6, 8,
;    TITLE OF INVENTION:  11, 16, 18, 31, 33 AND 56,
;    TITLE OF INVENTION:  USEFUL IN IMMUNOASSAY FOR
;    TITLE OF INVENTION:  DIAGNOSTIC PURPOSES
;    NUMBER OF SEQUENCES:  193
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  MASON & ASSOCIATES, P.A.
;      STREET:  17757 U.S. HWY. 19 NORTH, SUITE 500
;      CITY:  CLEARWATER
;      STATE:  FLORIDA
;      COUNTRY:  U.S.A.
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  Windows 3.0
;      SOFTWARE:  Microsoft Word 6.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/934,915
;      FILING DATE:  22-SEP-1997
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/949,836
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  LOUISE A. Foutch
;      REGISTRATION NUMBER:  37,133
;      REFERENCE/DOCKET NUMBER:  1946.6
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  813-538-3800
;      TELEFAX:  813-538-3820
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  69:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  20 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-934-915-69

  Query Match             66.4%;  Score 79;  DB 2;  Length 20;
  Best Local Similarity   70.6%;  Pred. No. 1.9e-05;
  Matches   12;  Conservative    2;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1 DICNTMHYTNWTHIYIC 17
           |: |||||||| ||:|
Db       4 DVHNTMHYTNWKFIYLC 20


RESULT 11
US-08-934-915-74
; Sequence 74, Application US/08934915
; Patent No. 5932412
;  GENERAL INFORMATION:
;    APPLICANT:  DILLNER, JOAKIM
;    APPLICANT:  DILLNER, LENA
;    APPLICANT:  CHENG, HWEE-MING
;    TITLE OF INVENTION:  SYNTHETIC PEPTIDES OF HUMAN
;    TITLE OF INVENTION:  PAPILLOMAVIRUS 1, 5, 6, 8,
;    TITLE OF INVENTION:  11, 16, 18, 31, 33 AND 56,
;    TITLE OF INVENTION:  USEFUL IN IMMUNOASSAY FOR
;    TITLE OF INVENTION:  DIAGNOSTIC PURPOSES
;    NUMBER OF SEQUENCES:  193
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  MASON & ASSOCIATES, P.A.
;      STREET:  17757 U.S. HWY. 19 NORTH, SUITE 500
;      CITY:  CLEARWATER
;      STATE:  FLORIDA
;      COUNTRY:  U.S.A.
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  Windows 3.0
;      SOFTWARE:  Microsoft Word 6.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/934,915
;      FILING DATE:  22-SEP-1997
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/949,836
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  LOUISE A. Foutch
;      REGISTRATION NUMBER:  37,133
;      REFERENCE/DOCKET NUMBER:  1946.6
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  813-538-3800
;      TELEFAX:  813-538-3820
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  74:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  20 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-934-915-74

  Query Match             47.9%;  Score 57;  DB 2;  Length 20;
  Best Local Similarity   76.9%;  Pred. No. 0.03;
  Matches   10;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       4 NTMHYTNWTHIYI 16
           ||| |||| |||
Db       7 NTMDYTNWGEIYI 19


RESULT 12
US-08-934-915-13
; Sequence 13, Application US/08934915
; Patent No. 5932412
;  GENERAL INFORMATION:
;    APPLICANT:  DILLNER, JOAKIM
;    APPLICANT:  DILLNER, LENA
;    APPLICANT:  CHENG, HWEE-MING
;    TITLE OF INVENTION:  SYNTHETIC PEPTIDES OF HUMAN
;    TITLE OF INVENTION:  PAPILLOMAVIRUS 1, 5, 6, 8,
;    TITLE OF INVENTION:  11, 16, 18, 31, 33 AND 56,
;    TITLE OF INVENTION:  USEFUL IN IMMUNOASSAY FOR
;    TITLE OF INVENTION:  DIAGNOSTIC PURPOSES
;    NUMBER OF SEQUENCES:  193
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  MASON & ASSOCIATES, P.A.
```

```
;       STREET:  17757 U.S. HWY. 19 NORTH, SUITE 500
;       CITY:  CLEARWATER
;       STATE:  FLORIDA
;       COUNTRY:  U.S.A.
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  Windows 3.0
;       SOFTWARE:  Microsoft Word 6.0
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/934,915
;       FILING DATE:  22-SEP-1997
;       CLASSIFICATION:  435
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  07/949,836
;       FILING DATE:
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  LOUISE A. Foutch
;       REGISTRATION NUMBER:  37,133
;       REFERENCE/DOCKET NUMBER:  1946.6
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  813-538-3800
;       TELEFAX:  813-538-3820
;       TELEX:
;   INFORMATION FOR SEQ ID NO:  13:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  20 amino acids
;       TYPE:  amino acid
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  peptide
US-08-934-915-13

  Query Match             43.7%;  Score 52;  DB 2;  Length 20;
  Best Local Similarity   61.5%;  Pred. No. 0.16;
  Matches    8;  Conservative    1;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       4 NTMHYTNWTHIYI 16
           | | |  |||||:
Db       7 NVMEYVVWTHIYL 19


RESULT 13
US-08-934-915-7
; Sequence 7, Application US/08934915
; Patent No. 5932412
;  GENERAL INFORMATION:
;    APPLICANT:  DILLNER, JOAKIM
;    APPLICANT:  DILLNER, LENA
;    APPLICANT:  CHENG, HWEE-MING
;    TITLE OF INVENTION:  SYNTHETIC PEPTIDES OF HUMAN
;    TITLE OF INVENTION:  PAPILLOMAVIRUS 1, 5, 6, 8,
;    TITLE OF INVENTION:  11, 16, 18, 31, 33 AND 56,
;    TITLE OF INVENTION:  USEFUL IN IMMUNOASSAY FOR
;    TITLE OF INVENTION:  DIAGNOSTIC PURPOSES
;    NUMBER OF SEQUENCES:  193
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  MASON & ASSOCIATES, P.A.
;      STREET:  17757 U.S. HWY. 19 NORTH, SUITE 500
;      CITY:  CLEARWATER
;      STATE:  FLORIDA
;      COUNTRY:  U.S.A.
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  Windows 3.0
;      SOFTWARE:  Microsoft Word 6.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/934,915
;      FILING DATE:  22-SEP-1997
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/949,836
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  LOUISE A. Foutch
;      REGISTRATION NUMBER:  37,133
;      REFERENCE/DOCKET NUMBER:  1946.6
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  813-538-3800
;      TELEFAX:  813-538-3820
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  7:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  20 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-934-915-7

  Query Match             39.5%;  Score 47;  DB 2;  Length 20;
  Best Local Similarity   53.8%;  Pred. No. 0.85;
  Matches    7;  Conservative    2;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       4 NTMHYTNWTHIYI 16
           ||| |  || :|:
Db       7 NTMDYVVWTDVYV 19


RESULT 14
US-09-000-094-46
; Sequence 46, Application US/09000094
; Patent No. 6365160
;   GENERAL INFORMATION:
;       APPLICANT: WEBB, Elizabeth Ann
;                  MARGETTS, Mary Brigid
;                  COX, John Cooper
;                  FRAZER, Ian
;                  MCMILLAN, Nigel Alan John
;                  WILLIAMS, Mark Philip
;                  MOLONEY, Margaret Bridget
;                  Holland
;                  EDWARDS, Stirling John
;       TITLE OF INVENTION: PAPILLOMAVIRUS POLYPROTEIN CONSTRUCTS
;       NUMBER OF SEQUENCES: 50
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: FOLEY & LARDNER
;            STREET: 3000 K Street, N.W.
;            CITY: Washington
;            STATE: D.C.
;            COUNTRY: U.S.A.
;            ZIP: 20007-5109
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Floppy disk
;            COMPUTER: IBM PC compatible
;            OPERATING SYSTEM: PC-DOS/MS-DOS
;            SOFTWARE: PatentIn Release #1.0, Version #1.30
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/09/000,094
;            FILING DATE: 21-Apr-1998
;            CLASSIFICATION: <Unknown>
;       PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: WO PCT/AU96/00473
;            FILING DATE: 26-JUL-1996
;            APPLICATION NUMBER: AU PN 4439/95
;            FILING DATE: 27-JUL-1995
;       ATTORNEY/AGENT INFORMATION:
;            NAME: BENT, Stephen A.
;            REGISTRATION NUMBER: 29,768
;            REFERENCE/DOCKET NUMBER: 017227/0137
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: (202) 672-5300
;            TELEFAX: (202) 672-5399
;   INFORMATION FOR SEQ ID NO: 46:
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 1587 amino acids
;            TYPE: amino acid
```

```
;           TOPOLOGY: linear
;       MOLECULE TYPE: protein
;       SEQUENCE DESCRIPTION: SEQ ID NO: 46:
US-09-000-094-46

  Query Match            39.5%;  Score 47;  DB 4;  Length 1587;
  Best Local Similarity  53.8%;  Pred. No. 97;
  Matches    7;  Conservative    2;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       4 NTMHYTNWTHIYI 16
           ||| |  || :|:
Db     165 NTMDYVVWTDVYV 177


RESULT 15
dimeric-glycoprotein-receptor-extracellular
; Sequence 2, Application US/08474986
; Patent No. 6372711
;   GENERAL INFORMATION:
;       APPLICANT: Kelton, Christie Ann
;                  Schweickhardt, Rene Lynn
;                  Cheng, Shirley Vui Yen
;                  Nugent, No. 6372711een Patrice
;       TITLE OF INVENTION: Human Follicle Stimulating
;                           Hormone Receptor
;       NUMBER OF SEQUENCES: 2
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: Stephan P. Williams,
;                       Ares-Serono, Inc.
;            STREET: Exchange Place, 37th floor
;            CITY: Boston
;            STATE: MA
;            COUNTRY: USA
;            ZIP: 02109
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: 3.5" diskette, 1.44 MB, high density
;            COMPUTER: IBM PS/2, model 55 SX
;            OPERATING SYSTEM: MS-DOS version 4.0
;            SOFTWARE: VAX/VMS Mass11 via Kermit to IBM MS-DOS
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/08/474,986
;            FILING DATE: 07-Jun-1995
;            CLASSIFICATION: <Unknown>
;       PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: 07/670,085
;            FILING DATE: <Unknown>
;       ATTORNEY/AGENT INFORMATION:
;            NAME: Williams, Stephan P.
;            REGISTRATION NUMBER: 28546
;            REFERENCE/DOCKET NUMBER: US/252
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: (617) 723-1300
;            TELEFAX: (617) 723-8923
;   INFORMATION FOR SEQ ID NO: 2:
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 695
;            TYPE: Amino acid
;            TOPOLOGY: Linear
;       MOLECULE TYPE: protein
;       FEATURE:
;            NAME/KEY:  signal sequence
;            LOCATION:  -17 to -1
;            IDENTIFICATION METHOD:  hydrophobic
;       FEATURE:
;            NAME/KEY:  putative amino-terminal extracellular domain
;            LOCATION:  1 to 349
;            IDENTIFICATION METHOD:  similarity with other
;
dimeric-glycoprotein-receptor-extracellular

  Query Match            37.0%;  Score 44;  DB 4;  Length 19;
  Best Local Similarity  50.0%;  Pred. No. 2.2;
  Matches    7;  Conservative    1;  Mismatches    6;  Indels    0;  Gaps    0;

Qy       4 NTMHYTNWTHIYIC 17
           |: ||    || ||
Db       6 NS,HYDROPHILIC 19


Search completed: January 14, 2003, 18:41:31
Job time : 15 secs
```



OM protein - protein search, using sw model

Run on: January 14, 2003, 18:39:58 ; Search time 11 Seconds
(without alignments)
33.511 Million cell updates/sec

Title: US-09-828-645-1
Perfect score: 119
Sequence: 1 DICNTMHYTNWTHIYICEE 19

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 118974 seqs, 19401057 residues

Total number of hits satisfying chosen parameters: 118974

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA:*
1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
7: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep:*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep:*
11: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep:*
13: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	52	43.7	334	10	US-09-886-055-365	Sequence 365, App
2	50	42.0	436	10	US-09-925-300-1640	Sequence 1640, Ap
3	43	36.1	87	10	US-09-925-301-1125	Sequence 1125, Ap
4	43	36.1	313	10	US-09-788-626-7	Sequence 7, Appli
5	43	36.1	314	10	US-09-886-055-141	Sequence 141, App
6	42.5	35.7	135	10	US-09-864-761-39255	Sequence 39255, A
7	42	35.3	162	10	US-09-747-155-211	Sequence 211, App
8	42	35.3	309	10	US-09-788-626-16	Sequence 16, Appl
9	42	35.3	340	10	US-09-886-055-465	Sequence 465, App
10	41	34.5	305	10	US-09-761-288-86	Sequence 86, Appl
11	41	34.5	311	10	US-09-886-055-419	Sequence 419, App
12	41	34.5	640	9	US-09-983-204-16	Sequence 16, Appl
13	41	34.5	640	9	US-10-133-157-5	Sequence 5, Appli
14	41	34.5	656	10	US-09-799-777-4	Sequence 4, Appli
15	40.5	34.0	52	10	US-09-864-761-44604	Sequence 44604, A
16	40.5	34.0	241	9	US-09-738-626-6844	Sequence 6844, Ap
17	40.5	34.0	310	10	US-09-886-055-485	Sequence 485, App
18	40	33.6	10	10	US-09-753-126-22	Sequence 22, Appl
19	40	33.6	10	10	US-09-753-126-139	Sequence 139, App
20	40	33.6	10	10	US-09-753-126-140	Sequence 140, App
21	40	33.6	216	10	US-09-747-155-49	Sequence 49, Appl
22	40	33.6	238	10	US-09-729-835-80	Sequence 80, Appl
23	40	33.6	309	10	US-09-886-055-383	Sequence 383, App
24	40	33.6	310	10	US-09-886-055-153	Sequence 153, App
25	40	33.6	333	10	US-09-771-209-71	Sequence 71, Appl
26	40	33.6	336	10	US-09-747-804-1	Sequence 1, Appli
27	40	33.6	419	12	US-10-005-202-2	Sequence 2, Appli
28	40	33.6	477	10	US-09-966-147-4	Sequence 4, Appli
29	40	33.6	525	10	US-09-871-503-2	Sequence 2, Appli
30	40	33.6	525	10	US-09-940-919-2	Sequence 2, Appli
31	40	33.6	822	10	US-09-966-147-2	Sequence 2, Appli
32	40	33.6	847	10	US-09-924-859A-5	Sequence 5, Appli
33	40	33.6	1350	10	US-09-893-238-17	Sequence 17, Appl
34	39.5	33.2	495	10	US-09-776-865-4	Sequence 4, Appli
35	39.5	33.2	495	10	US-09-915-181A-7	Sequence 7, Appli
36	39.5	33.2	536	10	US-09-776-865-2	Sequence 2, Appli
37	39	32.8	46	9	US-09-975-143-19	Sequence 19, Appl
38	39	32.8	160	10	US-09-970-711-38	Sequence 38, Appl
39	39	32.8	161	10	US-09-747-155-321	Sequence 321, App
40	39	32.8	161	10	US-09-747-155-323	Sequence 323, App
41	39	32.8	255	10	US-09-529-063-67	Sequence 67, Appl
42	39	32.8	281	10	US-09-864-761-43213	Sequence 43213, A
43	39	32.8	307	10	US-09-071-838-13	Sequence 13, Appl
44	39	32.8	338	10	US-09-886-055-427	Sequence 427, App
45	39	32.8	777	12	US-10-052-586-348	Sequence 348, App

ALIGNMENTS

RESULT 1
US-09-886-055-365
; Sequence 365, Application US/09886055
; Patent No. US20020132273A1
; GENERAL INFORMATION:
; APPLICANT: STRYER, LUBERT
; APPLICANT: ZOZULYA, SERGEY
; TITLE OF INVENTION: RECEPTOR FINGERPRINTING, SENSORY PERCEPTION, AND
; TITLE OF INVENTION: BIOSENSORS OF CHEMICAL SENSANTS
; FILE REFERENCE: 078003-0277150
; CURRENT APPLICATION NUMBER: US/09/886,055
; CURRENT FILING DATE: 2001-06-22
; PRIOR APPLICATION NUMBER: 60/213,812
; PRIOR FILING DATE: 2000-06-22
; NUMBER OF SEQ ID NOS: 522
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 365
; LENGTH: 334
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-886-055-365

Query Match 43.7%; Score 52; DB 10; Length 334;
Best Local Similarity 43.8%; Pred. No. 1.9;
Matches 7; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

```
Qy       2 ICNTMHYTNWTHIYIC 17
           :|  :|||   ||::|
Db     146 VCRPLHYTVLMHIHLC 161
```

RESULT 2
US-09-925-300-1640
; Sequence 1640, Application US/09925300
; Patent No. US20020151681A1
; GENERAL INFORMATION:
; APPLICANT: Craig Rosen,
; APPLICANT: Steve Ruben
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA101
; CURRENT APPLICATION NUMBER: US/09/925,300
; CURRENT FILING DATE: 2001-08-10

```
;  PRIOR APPLICATION NUMBER: PCT/US00/05988
;  PRIOR FILING DATE: 2000-03-08
;  PRIOR APPLICATION NUMBER: 60/124,270
;  PRIOR FILING DATE: 1999-03-12
;  NUMBER OF SEQ ID NOS: 1890
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1640
;   LENGTH: 436
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-925-300-1640

  Query Match             42.0%;  Score 50;  DB 10;  Length 436;
  Best Local Similarity   47.4%;  Pred. No. 4.7;
  Matches    9;  Conservative    2;  Mismatches    8;  Indels    0;  Gaps    0;

Qy      1 DICNTMHYTNWTHIYICEE 19
          ||    ||:||  : | ||
Db    299 DIFPARHYSNWETVLILEE 317


RESULT 3
US-09-925-301-1125
; Sequence 1125, Application US/09925301
; Patent No. US20020052308A1
; GENERAL INFORMATION:
;  APPLICANT: Rosen et al.
;  TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
;  FILE REFERENCE: PA106
;  CURRENT APPLICATION NUMBER: US/09/925,301
;  CURRENT FILING DATE:  2001-08-10
;  PRIOR APPLICATION NUMBER: PCT/US00/05882
;  PRIOR FILING DATE: 2000-03-08
;  PRIOR APPLICATION NUMBER: 60/124,270
;  PRIOR FILING DATE: 1999-03-12
;  NUMBER OF SEQ ID NOS: 1694
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1125
;   LENGTH: 87
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-925-301-1125

  Query Match             36.1%;  Score 43;  DB 10;  Length 87;
  Best Local Similarity   46.2%;  Pred. No. 9.8;
  Matches    6;  Conservative    3;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      3 CNTMHYTNWTHIY 15
          || | :  ||::|
Db     22 CNQMPFYMWTYLY 34


RESULT 4
US-09-788-626-7
; Sequence 7, Application US/09788626
; Patent No. US20020009762A1
; GENERAL INFORMATION:
;  APPLICANT: Flint, Andrew J.
;  APPLICANT:  Cool, Deborah E.
;  TITLE OF INVENTION: IMPROVED ASSAY FOR PROTEIN TYROSINE
;  TITLE OF INVENTION:  PHOSPHATES
;  FILE REFERENCE: 200125.401
;  CURRENT APPLICATION NUMBER: US/09/788,626
;  CURRENT FILING DATE:  2001-02-13
;  NUMBER OF SEQ ID NOS: 40
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 7
;   LENGTH: 313
;   TYPE: PRT
;   ORGANISM: Drosophila melanogaster
US-09-788-626-7

  Query Match             36.1%;  Score 43;  DB 10;  Length 313;
  Best Local Similarity   53.3%;  Pred. No. 34;
  Matches    8;  Conservative    0;  Mismatches    1;  Indels    6;  Gaps    1;

Qy      3 CNT------MHYTNW 11
          |||      |||||
Db    142 CNTEKLVYQYHYTNW 156


RESULT 5
US-09-886-055-141
; Sequence 141, Application US/09886055
; Patent No. US20020132273A1
; GENERAL INFORMATION:
;  APPLICANT: STRYER, LUBERT
;  APPLICANT:  ZOZULYA, SERGEY
;  TITLE OF INVENTION: RECEPTOR FINGERPRINTING, SENSORY PERCEPTION, AND
;  TITLE OF INVENTION:  BIOSENSORS OF CHEMICAL SENSANTS
;  FILE REFERENCE: 078003-0277150
;  CURRENT APPLICATION NUMBER: US/09/886,055
;  CURRENT FILING DATE:  2001-06-22
;  PRIOR APPLICATION NUMBER: 60/213,812
;  PRIOR FILING DATE: 2000-06-22
;  NUMBER OF SEQ ID NOS: 522
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 141
;   LENGTH: 314
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-886-055-141

  Query Match             36.1%;  Score 43;  DB 10;  Length 314;
  Best Local Similarity   37.5%;  Pred. No. 34;
  Matches    6;  Conservative    3;  Mismatches    7;  Indels    0;  Gaps    0;

Qy      2 ICNTMHYTNWTHIYIC 17
          :|| :|||      :|
Db    124 VCNPLHYTTTMTTRVC 139


RESULT 6
US-09-864-761-39255
; Sequence 39255, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
;  APPLICANT: Penn, Sharron G.
;  APPLICANT:  Rank, David R.
;  APPLICANT:  Hanzel, David K.
;  APPLICANT:  Chen, Wensheng
;  TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO
;  TITLE OF INVENTION:  GENE EXPRESSION ANALYSIS BY MICROARRAY
;  FILE REFERENCE: Aeomica-X-1
;  CURRENT APPLICATION NUMBER: US/09/864,761
;  CURRENT FILING DATE:  2001-05-23
;  PRIOR APPLICATION NUMBER: US 60/180,312
;  PRIOR FILING DATE: 2000-02-04
;  PRIOR APPLICATION NUMBER: US 60/207,456
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: US 09/632,366
;  PRIOR FILING DATE: 2000-08-03
;  PRIOR APPLICATION NUMBER: GB 24263.6
;  PRIOR FILING DATE: 2000-10-04
;  PRIOR APPLICATION NUMBER: US 60/236,359
;  PRIOR FILING DATE: 2000-09-27
;  PRIOR APPLICATION NUMBER: PCT/US01/00666
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00667
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00664
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00669
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00665
;  PRIOR FILING DATE: 2001-01-30
```

```
;  PRIOR APPLICATION NUMBER: PCT/US01/00668
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00663
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00662
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00661
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00670
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: US 60/234,687
;  PRIOR FILING DATE: 2000-09-21
;  PRIOR APPLICATION NUMBER: US 09/608,408
;  PRIOR FILING DATE: 2000-06-30
;  PRIOR APPLICATION NUMBER: US 09/774,203
;  PRIOR FILING DATE: 2001-01-29
;  NUMBER OF SEQ ID NOS: 49117
;  SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 39255
;   LENGTH: 135
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: MAP TO AC005815.1
;   OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 5.4
;   OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 4.2
;   OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 3.4
;   OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2.6
;   OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 2.8
;   OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 7.6
;   OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 4.2
;   OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 5.3
;   OTHER INFORMATION: EST_HUMAN HIT: BE739207.1, EVALUE 9.10e+00
;   OTHER INFORMATION: SWISSPROT HIT: P32588, EVALUE 1.60e+00
US-09-864-761-39255

  Query Match             35.7%;  Score 42.5;  DB 10;  Length 135;
  Best Local Similarity   38.9%;  Pred. No. 18;
  Matches    7;  Conservative    6;  Mismatches    4;  Indels    1;  Gaps    1;

Qy       1 DICNTMHYTNWTHIYICE 18
           :||  : :||  |:::||
Db      93 NICTIIPHTN-KHVHLCE 109


RESULT 7
US-09-747-155-211
; Sequence 211, Application US/09747155
; Patent No. US20020151692A1
; GENERAL INFORMATION:
;  APPLICANT: Rouquier, Sylvie
;  APPLICANT:  Giorgi, Dominique
;  TITLE OF INVENTION: No. US20020151692A1el Polypeptides and Nucleic Acids Encoding Sam
;  FILE REFERENCE: 19904-008 (C009B6834US)
;  CURRENT APPLICATION NUMBER: US/09/747,155
;  CURRENT FILING DATE:  2000-12-21
;  PRIOR APPLICATION NUMBER: 60/171,746
;  PRIOR FILING DATE: 1999-12-22
;  NUMBER OF SEQ ID NOS: 431
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 211
;   LENGTH: 162
;   TYPE: PRT
;   ORGANISM: Gorilla gorilla
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(487)
;   OTHER INFORMATION: Taxon = 9593; gene = GGO100; Accession DDBJ/EMBL/GenBank = AF1797
US-09-747-155-211

  Query Match             35.3%;  Score 42;  DB 10;  Length 162;
  Best Local Similarity   23.3%;  Pred. No. 25;
  Matches    7;  Conservative    6;  Mismatches    3;  Indels   14;  Gaps    1;

Qy       2 ICNTMHYT--------------NWTHIYIC 17
           ||: :|||              :|:  ::|
Db       3 ICHPLHYTFIMDQNTCIQLAVISWSSSFLC 32


RESULT 8
US-09-788-626-16
; Sequence 16, Application US/09788626
; Patent No. US20020009762A1
; GENERAL INFORMATION:
;  APPLICANT: Flint, Andrew J.
;  APPLICANT:  Cool, Deborah E.
;  TITLE OF INVENTION: IMPROVED ASSAY FOR PROTEIN TYROSINE
;  TITLE OF INVENTION:  PHOSPHATES
;  FILE REFERENCE: 200125.401
;  CURRENT APPLICATION NUMBER: US/09/788,626
;  CURRENT FILING DATE:  2001-02-13
;  NUMBER OF SEQ ID NOS: 40
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 16
;   LENGTH: 309
;   TYPE: PRT
;   ORGANISM: Drosophila melanogaster
US-09-788-626-16

  Query Match             35.3%;  Score 42;  DB 10;  Length 309;
  Best Local Similarity   37.5%;  Pred. No. 47;
  Matches    6;  Conservative    2;  Mismatches    8;  Indels    0;  Gaps    0;

Qy       2 ICNTMHYTNWTHIYIC 17
           | |  || :|    :|
Db     117 ILNDSHYADWVMFMLC 132


RESULT 9
US-09-886-055-465
; Sequence 465, Application US/09886055
; Patent No. US20020132273A1
; GENERAL INFORMATION:
;  APPLICANT: STRYER, LUBERT
;  APPLICANT:  ZOZULYA, SERGEY
;  TITLE OF INVENTION: RECEPTOR FINGERPRINTING, SENSORY PERCEPTION, AND
;  TITLE OF INVENTION:  BIOSENSORS OF CHEMICAL SENSANTS
;  FILE REFERENCE: 078003-0277150
;  CURRENT APPLICATION NUMBER: US/09/886,055
;  CURRENT FILING DATE:  2001-06-22
;  PRIOR APPLICATION NUMBER: 60/213,812
;  PRIOR FILING DATE: 2000-06-22
;  NUMBER OF SEQ ID NOS: 522
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 465
;   LENGTH: 340
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-886-055-465

  Query Match             35.3%;  Score 42;  DB 10;  Length 340;
  Best Local Similarity   43.8%;  Pred. No. 51;
  Matches    7;  Conservative    2;  Mismatches    7;  Indels    0;  Gaps    0;

Qy       2 ICNTMHYTNWTHIYIC 17
           ||  :|||   |  :|
Db     147 ICRPLHYTVIMHPQLC 162


RESULT 10
US-09-761-288-86
; Sequence 86, Application US/09761288
; Patent No. US20020065405A1
; GENERAL INFORMATION:
;  APPLICANT: Padigaru, Muralidhara
;  APPLICANT:  Prayaga, Sudhirdas
```

```
;  APPLICANT:  Taupier, Raymond J
;  APPLICANT:  Mishra, Vishnu
;  APPLICANT:  Tchernev, Velizar
;  APPLICANT:  Spytek, Kimberky
;  APPLICANT:  Li, Li
;  TITLE OF INVENTION: No. US20020065405A1el Polypeptides and Nucleic Acids Encoding Sam
;  FILE REFERENCE: 15966-638
;  CURRENT APPLICATION NUMBER: US/09/761,288
;  CURRENT FILING DATE:  2001-01-16
;  PRIOR APPLICATION NUMBER: 60/177,839
;  PRIOR FILING DATE: 2000-01-25
;  PRIOR APPLICATION NUMBER: 60/176,134
;  PRIOR FILING DATE: 2000-01-14
;  PRIOR APPLICATION NUMBER: 60/175,989
;  PRIOR FILING DATE: 2000-01-13
;  PRIOR APPLICATION NUMBER: 60/218,324
;  PRIOR FILING DATE: 2000-07-14
;  PRIOR APPLICATION NUMBER: 60/220,253
;  PRIOR FILING DATE: 2000-07-24
;  PRIOR APPLICATION NUMBER: 60/178,191
;  PRIOR FILING DATE: 2000-01-26
;  PRIOR APPLICATION NUMBER: 60/178,227
;  PRIOR FILING DATE: 2000-01-26
;  PRIOR APPLICATION NUMBER: 60/220,590
;  PRIOR FILING DATE: 2000-07-25
;  NUMBER OF SEQ ID NOS: 95
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 86
;   LENGTH: 305
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-761-288-86

  Query Match             34.5%;  Score 41;  DB 10;  Length 305;
  Best Local Similarity   37.5%;  Pred. No. 64;
  Matches    6;  Conservative    5;  Mismatches    5;  Indels    0;  Gaps    0;

Qy      2 ICNTMHYTNWTHIYIC 17
          ||  :||:|  :: :|
Db    126 ICIPLHYSNVMNMKVC 141


RESULT 11
US-09-886-055-419
; Sequence 419, Application US/09886055
; Patent No. US20020132273A1
; GENERAL INFORMATION:
;  APPLICANT: STRYER, LUBERT
;  APPLICANT:  ZOZULYA, SERGEY
;  TITLE OF INVENTION: RECEPTOR FINGERPRINTING, SENSORY PERCEPTION, AND
;  TITLE OF INVENTION:  BIOSENSORS OF CHEMICAL SENSANTS
;  FILE REFERENCE: 078003-0277150
;  CURRENT APPLICATION NUMBER: US/09/886,055
;  CURRENT FILING DATE:  2001-06-22
;  PRIOR APPLICATION NUMBER: 60/213,812
;  PRIOR FILING DATE: 2000-06-22
;  NUMBER OF SEQ ID NOS: 522
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 419
;   LENGTH: 311
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-886-055-419

  Query Match             34.5%;  Score 41;  DB 10;  Length 311;
  Best Local Similarity   29.4%;  Pred. No. 65;
  Matches    5;  Conservative    5;  Mismatches    7;  Indels    0;  Gaps    0;

Qy      2 ICNTMHYTNWTHIYICE 18
          :|  :||    | ::|:
Db    126 VCRPLHYMAIMHPHLCQ 142


RESULT 12
US-09-983-204-16
; Sequence 16, Application US/09983204
; Patent No. US20020173000A1
; GENERAL INFORMATION:
;  APPLICANT: RENARD, STEPHANE
;  APPLICANT:  BESNARD, FRANCOIS
;  APPLICANT:  GRAHAM, DAVID
;  TITLE OF INVENTION: SODIUM CHANNEL RECEPTOR
;  FILE REFERENCE: 07586.0010
;  CURRENT APPLICATION NUMBER: US/09/983,204
;  CURRENT FILING DATE:  2001-10-23
;  PRIOR APPLICATION NUMBER: 09/424,666
;  PRIOR FILING DATE: 2001-02-22
;  PRIOR APPLICATION NUMBER: PCT/EP98/02884
;  PRIOR FILING DATE: 1998-05-15
;  PRIOR APPLICATION NUMBER: 97401196.7
;  PRIOR FILING DATE: 1997-05-30
;  NUMBER OF SEQ ID NOS: 19
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 16
;   LENGTH: 640
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: HNACHB
US-09-983-204-16

  Query Match             34.5%;  Score 41;  DB 9;  Length 640;
  Best Local Similarity   38.5%;  Pred. No. 1.3e+02;
  Matches    5;  Conservative    2;  Mismatches    6;  Indels    0;  Gaps    0;

Qy      3 CNTMHYTNWTHIY 15
          ||   : :| | |
Db    415 CNNRDFPDWAHCY 427


RESULT 13
US-10-133-157-5
; Sequence 5, Application US/10133157
; Publication No. US20020184054A1
; GENERAL INFORMATION:
;  APPLICANT: CALLAMARAS, NICHOLAS
;  APPLICANT:  CHANG, HONG
;  TITLE OF INVENTION: HIGH THROUGHPUT CELL-BASED ASSAY FOR MONITORING SODIUM
;  TITLE OF INVENTION:  CHANNEL ACTIVITY AND DISCOVERY OF SALTY TASTE
;  TITLE OF INVENTION:  MODULATING COMPOUNDS
;  FILE REFERENCE: 078003-0280790
;  CURRENT APPLICATION NUMBER: US/10/133,157
;  CURRENT FILING DATE:  2002-04-26
;  PRIOR APPLICATION NUMBER: 60/287,413
;  PRIOR FILING DATE: 2001-05-01
;  NUMBER OF SEQ ID NOS: 14
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
;   LENGTH: 640
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-133-157-5

  Query Match             34.5%;  Score 41;  DB 9;  Length 640;
  Best Local Similarity   38.5%;  Pred. No. 1.3e+02;
  Matches    5;  Conservative    2;  Mismatches    6;  Indels    0;  Gaps    0;

Qy      3 CNTMHYTNWTHIY 15
          ||   : :| | |
Db    415 CNNRDFPDWAHCY 427


RESULT 14
US-09-799-777-4
; Sequence 4, Application US/09799777
; Patent No. US20020091244A1
```

```
;   GENERAL INFORMATION:
;        APPLICANT: Lal, Preeti
;                   Hillman, Jennifer L.
;                   Corley, Neil C.
;                   Guegler, Karl J.
;                   Baugh, Mariah
;                   Sather, Susan
;                   Shah, Purvi
;        TITLE OF INVENTION: HUMAN SIGNAL PEPTIDE-CONTAINING PROTEINS
;        NUMBER OF SEQUENCES: 154
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
;             STREET: 3174 PORTER DRIVE
;             CITY: PALO ALTO
;             STATE: CALIFORNIA
;             COUNTRY: USA
;             ZIP: 94304
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/09/799,777
;             FILING DATE: 06-Mar-2001
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: US/09/002,485
;             FILING DATE: <Unknown>
;        ATTORNEY/AGENT INFORMATION:
;             NAME: BILLINGS, LUCY J.
;             REGISTRATION NUMBER: 36,749
;             REFERENCE/DOCKET NUMBER: PF-0459 US
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: (650) 855-0555
;             TELEFAX: (650) 845-4166
;   INFORMATION FOR SEQ ID NO: 4:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 656 amino acids
;             TYPE: amino acid
;             STRANDEDNESS: single
;             TOPOLOGY: linear
;        IMMEDIATE SOURCE:
;             LIBRARY: SYNORAT03
;             CLONE: 693453
;        SEQUENCE DESCRIPTION: SEQ ID NO: 4 :
US-09-799-777-4

  Query Match             34.5%;  Score 41;  DB 10;  Length 656;
  Best Local Similarity   30.8%;  Pred. No. 1.4e+02;
  Matches    4;  Conservative    5;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       6 MHYTNWTHIYICE 18
           :|  :|  :::||
Db     164 LHLLDWVRLHVCE 176


RESULT 15
US-09-864-761-44604
; Sequence 44604, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
;  APPLICANT: Penn, Sharron G.
;  APPLICANT:  Rank, David R.
;  APPLICANT:  Hanzel, David K.
;  APPLICANT:  Chen, Wensheng
;  TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
;  TITLE OF INVENTION:  GENE EXPRESSION ANALYSIS BY MICROARRAY
;  FILE REFERENCE: Aeomica-X-1
;  CURRENT APPLICATION NUMBER: US/09/864,761
;  CURRENT FILING DATE:  2001-05-23
;  PRIOR APPLICATION NUMBER: US 60/180,312
;  PRIOR FILING DATE: 2000-02-04
;  PRIOR APPLICATION NUMBER: US 60/207,456
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: US 09/632,366
;  PRIOR FILING DATE: 2000-08-03
;  PRIOR APPLICATION NUMBER: GB 24263.6
;  PRIOR FILING DATE: 2000-10-04
;  PRIOR APPLICATION NUMBER: US 60/236,359
;  PRIOR FILING DATE: 2000-09-27
;  PRIOR APPLICATION NUMBER: PCT/US01/00666
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00667
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00664
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00669
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00665
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00668
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00663
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00662
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00661
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00670
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: US 60/234,687
;  PRIOR FILING DATE: 2000-09-21
;  PRIOR APPLICATION NUMBER: US 09/608,408
;  PRIOR FILING DATE: 2000-06-30
;  PRIOR APPLICATION NUMBER: US 09/774,203
;  PRIOR FILING DATE: 2001-01-29
;  NUMBER OF SEQ ID NOS: 49117
;  SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 44604
;   LENGTH: 52
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: MAP TO AF241727.1
;   OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 0.55
;   OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.47
;   OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.75
;   OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 7.2
;   OTHER INFORMATION: SWISSPROT HIT: P41732, EVALUE 1.00e-28
;   OTHER INFORMATION: EST_HUMAN HIT: BE741256.1, EVALUE 2.00e-27
US-09-864-761-44604

  Query Match             34.0%;  Score 40.5;  DB 10;  Length 52;
  Best Local Similarity   47.1%;  Pred. No. 13;
  Matches    8;  Conservative    1;  Mismatches    7;  Indels    1;  Gaps    1;

Qy       3 CNTMHYTNW-THIYICE 18
           |   :|||| |  |  |
Db       4 CGVQNYTNWSTSPYFLE 20


Search completed: January 14, 2003, 18:41:49
Job time : 12 secs
```

GenCore version 5.1.3

OM protein - protein search, using sw model

Run on: January 14, 2003, 18:38:33 ; Search time 29 Seconds
(without alignments)
134.996 Million cell updates/sec

Title: US-09-828-645-1
Perfect score: 119
Sequence: 1 DICNTMHYTNWTHIYICEE 19

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_21:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	119	100.0	359	12	Q918T0	Q918t0 human papil
2	119	100.0	361	12	Q918U7	Q918u7 human papil
3	119	100.0	361	12	Q918U5	Q918u5 human papil
4	119	100.0	361	12	Q918U3	Q918u3 human papil
5	119	100.0	361	12	Q918U1	Q918u1 human papil
6	119	100.0	365	12	Q9E842	Q9e842 human papil
7	119	100.0	365	12	Q9E841	Q9e841 human papil
8	119	100.0	365	12	Q9E839	Q9e839 human papil
9	119	100.0	365	12	Q9YV76	Q9yv76 human papil
10	119	100.0	365	12	Q9YV75	Q9yv75 human papil
11	119	100.0	365	12	Q9YV74	Q9yv74 human papil
12	119	100.0	365	12	Q9YIV0	Q9yiv0 human papil
13	119	100.0	365	12	Q9YV73	Q9yv73 human papil
14	113	95.0	356	12	Q918T5	Q918t5 human papil
15	113	95.0	356	12	Q918T3	Q918t3 human papil
16	113	95.0	361	12	Q918T8	Q918t8 human papil
17	110	92.4	365	12	Q9YV72	Q9yv72 human papil
18	107	89.9	365	12	Q9E840	Q9e840 human papil
19	107	89.9	365	12	Q9WQX4	Q9wqx4 human papil
20	64.5	54.2	370	12	Q81967	Q81967 human papil
21	64	53.8	314	5	Q9NE71	Q9ne71 leishmania
22	62	52.1	377	12	Q9WHG3	Q9whg3 human papil
23	60	50.4	359	12	Q9IR56	Q9ir56 human papil
24	60	50.4	359	12	Q993Z7	Q993z7 human papil
25	59	49.6	380	12	Q9WNM7	Q9wnm7 human papil
26	57	47.9	368	12	O90726	O90726 human papil
27	54	45.4	376	12	O37387	O37387 common chim
28	53	44.5	469	12	O56937	O56937 human papil
29	51	42.9	350	12	Q82008	Q82008 human papil
30	50	42.0	278	4	Q9UJ05	Q9uj05 homo sapien
31	50	42.0	284	4	Q96BV3	Q96bv3 homo sapien
32	50	42.0	478	12	O56943	O56943 human papil
33	48	40.3	587	5	Q960P7	Q960p7 drosophila
34	47.5	39.9	702	15	Q9QC07	Q9qc07 human endog
35	47	39.5	368	12	Q9YYP1	Q9yyp1 human papil
36	47	39.5	444	12	O56955	O56955 human papil
37	47	39.5	453	10	Q9LU91	Q9lu91 arabidopsis
38	47	39.5	776	12	P88981	P88981 murid herpe
39	46.5	39.1	519	10	Q9LFX7	Q9lfx7 arabidopsis
40	46.5	39.1	676	16	Q8XQP3	Q8xqp3 ralstonia s
41	46	38.7	227	15	O11515	O11515 human immun
42	46	38.7	288	10	Q9AVH4	Q9avh4 pisum sativ
43	46	38.7	311	11	Q8VGX0	Q8vgx0 mus musculu
44	46	38.7	311	11	Q8VFA7	Q8vfa7 mus musculu
45	46	38.7	311	11	Q8VFA6	Q8vfa6 mus musculu

ALIGNMENTS

RESULT 1
Q918T0
ID Q918T0 PRELIMINARY; PRT; 359 AA.
AC Q918T0;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-JUN-2002, (TrEMBLrel. 21, Last annotation update)
DE E2 protein (Fragment).
GN E2.
OS Human papillomavirus type 16.
OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC Papillomavirus.
OX NCBI_TaxID=10581;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HPV16E2CC8;
RA Watts K.J., Thompson C.H., Cossart Y.E., Rose B.R.;
RT "Sequence variation and physical state of human Papillomavirus type 16
RT cervical cancer isolates from Australia and New Caledonia.";
RL Submitted (AUG-2001) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AF407221; AAL01407.1; -.
DR InterPro; IPR000427; E2_C.
DR InterPro; IPR001866; E2_N.
DR Pfam; PF00511; E2_C; 1.
DR Pfam; PF00508; E2_N; 1.
DR ProDom; PD000672; E2_C; 1.
DR ProDom; PD000678; E2_N; 1.
FT NON_TER 1 1
FT NON_TER 359 359
SQ SEQUENCE 359 AA; 41122 MW; 375DC218679CEA89 CRC64;

Query Match 100.0%; Score 119; DB 12; Length 359;
Best Local Similarity 100.0%; Pred. No. 8.6e-11;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DICNTMHYTNWTHIYICEE 19
||||||||||||||||||||
Db 121 DICNTMHYTNWTHIYICEE 139

```
RESULT 2
Q918U7
ID   Q918U7      PRELIMINARY;      PRT;   361 AA.
AC   Q918U7;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   E2 protein (Fragment).
GN   E2.
OS   Human papillomavirus type 16.
OC   Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC   Papillomavirus.
OX   NCBI_TaxID=10581;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=HPV16E2CC1;
RA   Watts K.J., Thompson C.H., Cossart Y.E., Rose B.R.;
RT   "Sequence variation and physical state of human Papillomavirus type 16
RT   cervical cancer isolates from Australia and New Caledonia.";
RL   Submitted (AUG-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF407214; AAL01386.1; -.
DR   InterPro; IPR000427; E2_C.
DR   InterPro; IPR001866; E2_N.
DR   Pfam; PF00511; E2_C; 1.
DR   Pfam; PF00508; E2_N; 1.
DR   ProDom; PD000672; E2_C; 1.
DR   ProDom; PD000678; E2_N; 1.
FT   NON_TER       1       1
FT   NON_TER     361     361
SQ   SEQUENCE   361 AA;  41335 MW;  87D0B9CEBCA9558D CRC64;

  Query Match             100.0%;  Score 119;  DB 12;  Length 361;
  Best Local Similarity   100.0%;  Pred. No. 8.7e-11;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 DICNTMHYTNWTHIYICEE 19
          ||||||||||||||||||||
Db    121 DICNTMHYTNWTHIYICEE 139


RESULT 3
Q918U5
ID   Q918U5      PRELIMINARY;      PRT;   361 AA.
AC   Q918U5;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   E2 protein (Fragment).
GN   E2.
OS   Human papillomavirus type 16.
OC   Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC   Papillomavirus.
OX   NCBI_TaxID=10581;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=HPV16E2CC2;
RA   Watts K.J., Thompson C.H., Cossart Y.E., Rose B.R.;
RT   "Sequence variation and physical state of human Papillomavirus type 16
RT   cervical cancer isolates from Australia and New Caledonia.";
RL   Submitted (AUG-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF407215; AAL01389.1; -.
DR   InterPro; IPR000427; E2_C.
DR   InterPro; IPR001866; E2_N.
DR   Pfam; PF00511; E2_C; 1.
DR   Pfam; PF00508; E2_N; 1.
DR   ProDom; PD000672; E2_C; 1.
DR   ProDom; PD000678; E2_N; 1.
FT   NON_TER       1       1
FT   NON_TER     361     361
SQ   SEQUENCE   361 AA;  41395 MW;  9CD109CEBCA9558D CRC64;

  Query Match             100.0%;  Score 119;  DB 12;  Length 361;
  Best Local Similarity   100.0%;  Pred. No. 8.7e-11;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 DICNTMHYTNWTHIYICEE 19
          ||||||||||||||||||||
Db    121 DICNTMHYTNWTHIYICEE 139


RESULT 4
Q918U3
ID   Q918U3      PRELIMINARY;      PRT;   361 AA.
AC   Q918U3;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   E2 protein (Fragment).
GN   E2.
OS   Human papillomavirus type 16.
OC   Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC   Papillomavirus.
OX   NCBI_TaxID=10581;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=HPV16E2CC3;
RA   Watts K.J., Thompson C.H., Cossart Y.E., Rose B.R.;
RT   "Sequence variation and physical state of human Papillomavirus type 16
RT   cervical cancer isolates from Australia and New Caledonia.";
RL   Submitted (AUG-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF407216; AAL01392.1; -.
DR   InterPro; IPR000427; E2_C.
DR   InterPro; IPR001866; E2_N.
DR   Pfam; PF00511; E2_C; 1.
DR   Pfam; PF00508; E2_N; 1.
DR   ProDom; PD000672; E2_C; 1.
DR   ProDom; PD000678; E2_N; 1.
FT   NON_TER       1       1
FT   NON_TER     361     361
SQ   SEQUENCE   361 AA;  41356 MW;  B973EA921D39BD97 CRC64;

  Query Match             100.0%;  Score 119;  DB 12;  Length 361;
  Best Local Similarity   100.0%;  Pred. No. 8.7e-11;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 DICNTMHYTNWTHIYICEE 19
          ||||||||||||||||||||
Db    121 DICNTMHYTNWTHIYICEE 139


RESULT 5
Q918U1
ID   Q918U1      PRELIMINARY;      PRT;   361 AA.
AC   Q918U1;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   E2 protein (Fragment).
GN   E2.
OS   Human papillomavirus type 16.
OC   Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC   Papillomavirus.
OX   NCBI_TaxID=10581;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=HPV16E2CC4;
RA   Watts K.J., Thompson C.H., Cossart Y.E., Rose B.R.;
RT   "Sequence variation and physical state of human Papillomavirus type 16
RT   cervical cancer isolates from Australia and New Caledonia.";
RL   Submitted (AUG-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF407217; AAL01395.1; -.
DR   InterPro; IPR000427; E2_C.
DR   InterPro; IPR001866; E2_N.
DR   Pfam; PF00511; E2_C; 1.
DR   Pfam; PF00508; E2_N; 1.
```

```
DR   ProDom; PD000672; E2_C; 1.
DR   ProDom; PD000678; E2_N; 1.
FT   NON_TER       1      1
FT   NON_TER     361    361
SQ   SEQUENCE   361 AA;  41394 MW;  E5E738AB8BAEF4C9 CRC64;

  Query Match             100.0%;  Score 119;  DB 12;  Length 361;
  Best Local Similarity   100.0%;  Pred. No. 8.7e-11;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 DICNTMHYTNWTHIYICEE 19
           |||||||||||||||||||
Db     121 DICNTMHYTNWTHIYICEE 139


RESULT 6
Q9E842
ID   Q9E842      PRELIMINARY;      PRT;   365 AA.
AC   Q9E842;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   E2 protein variant.
GN   E2.
OS   Human papillomavirus type 16.
OC   Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC   Papillomavirus.
OX   NCBI_TaxID=10581;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=DAG1;
RX   MEDLINE=20495818; PubMed=11040943;
RA   Graham D.A., Herrington C.S.;
RT   "HPV-16 E2 gene disruption and sequence variation in CIN 3 lesions and
RT   invasive squamous cell carcinomas of the cervix: relation to numerical
RT   chromosome abnormalities.";
RL   Mol. Pathol. 53:201-206(2000).
DR   EMBL; AF193425; AAG31352.1; -.
DR   HSSP; P17383; 1DHM.
DR   InterPro; IPR000427; E2_C.
DR   InterPro; IPR001866; E2_N.
DR   Pfam; PF00511; E2_C; 1.
DR   Pfam; PF00508; E2_N; 1.
DR   ProDom; PD000672; E2_C; 1.
DR   ProDom; PD000678; E2_N; 1.
SQ   SEQUENCE   365 AA;  41805 MW;  A647FDDDEE056C0D CRC64;

  Query Match             100.0%;  Score 119;  DB 12;  Length 365;
  Best Local Similarity   100.0%;  Pred. No. 8.8e-11;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 DICNTMHYTNWTHIYICEE 19
           |||||||||||||||||||
Db     124 DICNTMHYTNWTHIYICEE 142


RESULT 7
Q9E841
ID   Q9E841      PRELIMINARY;      PRT;   365 AA.
AC   Q9E841;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   E2 protein variant.
GN   E2.
OS   Human papillomavirus type 16.
OC   Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC   Papillomavirus.
OX   NCBI_TaxID=10581;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=DAG2;
RX   MEDLINE=20495818; PubMed=11040943;
RA   Graham D.A., Herrington C.S.;
RT   "HPV-16 E2 gene disruption and sequence variation in CIN 3 lesions and
RT   invasive squamous cell carcinomas of the cervix: relation to numerical
RT   chromosome abnormalities.";
RL   Mol. Pathol. 53:201-206(2000).
DR   EMBL; AF193426; AAG31353.1; -.
DR   HSSP; P17383; 1DHM.
DR   InterPro; IPR000427; E2_C.
DR   InterPro; IPR001866; E2_N.
DR   Pfam; PF00511; E2_C; 1.
DR   Pfam; PF00508; E2_N; 1.
DR   ProDom; PD000672; E2_C; 1.
DR   ProDom; PD000678; E2_N; 1.
SQ   SEQUENCE   365 AA;  41801 MW;  24682DDF132B7543 CRC64;

  Query Match             100.0%;  Score 119;  DB 12;  Length 365;
  Best Local Similarity   100.0%;  Pred. No. 8.8e-11;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 DICNTMHYTNWTHIYICEE 19
           |||||||||||||||||||
Db     124 DICNTMHYTNWTHIYICEE 142


RESULT 8
Q9E839
ID   Q9E839      PRELIMINARY;      PRT;   365 AA.
AC   Q9E839;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   E2 protein variant.
GN   E2.
OS   Human papillomavirus type 16.
OC   Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC   Papillomavirus.
OX   NCBI_TaxID=10581;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=DAG6;
RX   MEDLINE=20495818; PubMed=11040943;
RA   Graham D.A., Herrington C.S.;
RT   "HPV-16 E2 gene disruption and sequence variation in CIN 3 lesions and
RT   invasive squamous cell carcinomas of the cervix: relation to numerical
RT   chromosome abnormalities.";
RL   Mol. Pathol. 53:201-206(2000).
DR   EMBL; AF193430; AAG31357.1; -.
DR   HSSP; P17383; 1DHM.
DR   InterPro; IPR000427; E2_C.
DR   InterPro; IPR001866; E2_N.
DR   Pfam; PF00511; E2_C; 1.
DR   Pfam; PF00508; E2_N; 1.
DR   ProDom; PD000672; E2_C; 1.
DR   ProDom; PD000678; E2_N; 1.
SQ   SEQUENCE   365 AA;  41809 MW;  3EA67F918055C020 CRC64;

  Query Match             100.0%;  Score 119;  DB 12;  Length 365;
  Best Local Similarity   100.0%;  Pred. No. 8.8e-11;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 DICNTMHYTNWTHIYICEE 19
           |||||||||||||||||||
Db     124 DICNTMHYTNWTHIYICEE 142


RESULT 9
Q9YV76
ID   Q9YV76      PRELIMINARY;      PRT;   365 AA.
AC   Q9YV76;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Regulatory protein E2.
```

```
GN   E2.
OS   Human papillomavirus type 16.
OC   Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC   Papillomavirus.
OX   NCBI_TaxID=10581;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=H990;
RX   MEDLINE=99226956; PubMed=10211974;
RA   Veress G., Szarka K., Dong X.P., Gergely L., Pfister H.;
RT   "Functional significance of sequence variation in the E2 gene and the
RT   long control region of human papillomavirus type 16.";
RL   J. Gen. Virol. 80:1035-1043(1999).
DR   EMBL; AF067031; AAD03827.1; -.
DR   HSSP; P17383; 1DHM.
DR   InterPro; IPR000427; E2_C.
DR   InterPro; IPR001866; E2_N.
DR   Pfam; PF00511; E2_C; 1.
DR   Pfam; PF00508; E2_N; 1.
DR   ProDom; PD000672; E2_C; 1.
DR   ProDom; PD000678; E2_N; 1.
SQ   SEQUENCE   365 AA;  41841 MW;  2745A882B246E112 CRC64;

  Query Match             100.0%;  Score 119;  DB 12;  Length 365;
  Best Local Similarity   100.0%;  Pred. No. 8.8e-11;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 DICNTMHYTNWTHIYICEE 19
           |||||||||||||||||||
Db     124 DICNTMHYTNWTHIYICEE 142


RESULT 10
Q9YV75
ID   Q9YV75      PRELIMINARY;      PRT;   365 AA.
AC   Q9YV75;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Regulatory protein E2.
GN   E2.
OS   Human papillomavirus type 16.
OC   Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC   Papillomavirus.
OX   NCBI_TaxID=10581;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=H955;
RX   MEDLINE=99226956; PubMed=10211974;
RA   Veress G., Szarka K., Dong X.P., Gergely L., Pfister H.;
RT   "Functional significance of sequence variation in the E2 gene and the
RT   long control region of human papillomavirus type 16.";
RL   J. Gen. Virol. 80:1035-1043(1999).
DR   EMBL; AF067032; AAD03828.1; -.
DR   HSSP; P17383; 1DHM.
DR   InterPro; IPR000427; E2_C.
DR   InterPro; IPR001866; E2_N.
DR   Pfam; PF00511; E2_C; 1.
DR   Pfam; PF00508; E2_N; 1.
DR   ProDom; PD000672; E2_C; 1.
DR   ProDom; PD000678; E2_N; 1.
SQ   SEQUENCE   365 AA;  41816 MW;  3EB2EDDF132B7112 CRC64;

  Query Match             100.0%;  Score 119;  DB 12;  Length 365;
  Best Local Similarity   100.0%;  Pred. No. 8.8e-11;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 DICNTMHYTNWTHIYICEE 19
           |||||||||||||||||||
Db     124 DICNTMHYTNWTHIYICEE 142

[redacted]ULT 11
```

```
Q9YV74
ID   Q9YV74      PRELIMINARY;      PRT;   365 AA.
AC   Q9YV74;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Regulatory protein E2.
GN   E2.
OS   Human papillomavirus type 16.
OC   Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC   Papillomavirus.
OX   NCBI_TaxID=10581;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=H982;
RX   MEDLINE=99226956; PubMed=10211974;
RA   Veress G., Szarka K., Dong X.P., Gergely L., Pfister H.;
RT   "Functional significance of sequence variation in the E2 gene and the
RT   long control region of human papillomavirus type 16.";
RL   J. Gen. Virol. 80:1035-1043(1999).
DR   EMBL; AF067033; AAD03829.1; -.
DR   HSSP; P17383; 1DHM.
DR   InterPro; IPR000427; E2_C.
DR   InterPro; IPR001866; E2_N.
DR   Pfam; PF00511; E2_C; 1.
DR   Pfam; PF00508; E2_N; 1.
DR   ProDom; PD000672; E2_C; 1.
DR   ProDom; PD000678; E2_N; 1.
SQ   SEQUENCE   365 AA;  41831 MW;  17B9C0C9470D3376 CRC64;

  Query Match             100.0%;  Score 119;  DB 12;  Length 365;
  Best Local Similarity   100.0%;  Pred. No. 8.8e-11;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 DICNTMHYTNWTHIYICEE 19
           |||||||||||||||||||
Db     124 DICNTMHYTNWTHIYICEE 142


RESULT 12
Q9YIV0
ID   Q9YIV0      PRELIMINARY;      PRT;   365 AA.
AC   Q9YIV0;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Regulatory protein E2 (E2 protein variant).
GN   E2.
OS   Human papillomavirus type 16.
OC   Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC   Papillomavirus.
OX   NCBI_TaxID=10581;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=H050, AND H936;
RA   Veress G., Szarka K., Dong X.-P., Gergely L., Pfister H.;
RT   "Sequence variation in the E2 gene and in the long control region of
RT   human papillomavirus type 16.";
RL   Submitted (MAY-1998) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=16W12E;
RA   Flores E.R., Nelson J.H., Lambert P.F.;
RT   "Establishment of the human papillomavirus life cycle in an
RT   immortalized human foreskin keratinocyte cell line.";
RL   Submitted (FEB-1999) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=DAG3;
RX   MEDLINE=20495818; PubMed=11040943;
RA   Graham D.A., Herrington C.S.;
RT   "HPV-16 E2 gene disruption and sequence variation in CIN 3 lesions and
RT   invasive squamous cell carcinomas of the cervix: relation to numerical
```

```
RT   chromosome abnormalities.";
RL   Mol. Pathol. 53:201-206(2000).
DR   EMBL; AF067030; AAD03826.1; -.
DR   EMBL; AF067029; AAD03825.1; -.
DR   EMBL; AF125673; AAD33255.1; -.
DR   EMBL; AF193427; AAG31354.1; -.
DR   HSSP; P17383; 1DHM.
DR   InterPro; IPR000427; E2_C.
DR   InterPro; IPR001866; E2_N.
DR   Pfam; PF00511; E2_C; 1.
DR   Pfam; PF00508; E2_N; 1.
DR   ProDom; PD000672; E2_C; 1.
DR   ProDom; PD000678; E2_N; 1.
SQ   SEQUENCE   365 AA;  41815 MW;  24682DDF132B7112 CRC64;

  Query Match             100.0%;  Score 119;  DB 12;  Length 365;
  Best Local Similarity   100.0%;  Pred. No. 8.8e-11;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 DICNTMHYTNWTHIYICEE 19
          ||||||||||||||||||||
Db    124 DICNTMHYTNWTHIYICEE 142


RESULT 13
Q9YV73
ID   Q9YV73      PRELIMINARY;      PRT;   365 AA.
AC   Q9YV73;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Regulatory protein E2 (E2 protein variant).
GN   E2.
OS   Human papillomavirus type 16.
OC   Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC   Papillomavirus.
OX   NCBI_TaxID=10581;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=H059;
RA   Veress G., Szarka K., Dong X.-P., Gergely L., Pfister H.;
RT   "Sequence variation in the E2 gene and in the long control region of
RT   human papillomavirus type 16.";
RL   Submitted (MAY-1998) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=DAG4;
RX   MEDLINE=20495818; PubMed=11040943;
RA   Graham D.A., Herrington C.S.;
RT   "HPV-16 E2 gene disruption and sequence variation in CIN 3 lesions and
RT   invasive squamous cell carcinomas of the cervix: relation to numerical
RT   chromosome abnormalities.";
RL   Mol. Pathol. 53:201-206(2000).
DR   EMBL; AF067034; AAD03830.1; -.
DR   EMBL; AF193428; AAG31355.1; -.
DR   HSSP; P17383; 1DHM.
DR   InterPro; IPR000427; E2_C.
DR   InterPro; IPR001866; E2_N.
DR   Pfam; PF00511; E2_C; 1.
DR   Pfam; PF00508; E2_N; 1.
DR   ProDom; PD000672; E2_C; 1.
DR   ProDom; PD000678; E2_N; 1.
SQ   SEQUENCE   365 AA;  41830 MW;  2B78DCC2DCF450C8 CRC64;

  Query Match             100.0%;  Score 119;  DB 12;  Length 365;
  Best Local Similarity   100.0%;  Pred. No. 8.8e-11;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 DICNTMHYTNWTHIYICEE 19
          ||||||||||||||||||||
Db    124 DICNTMHYTNWTHIYICEE 142


RESULT 14
Q918T5
ID   Q918T5      PRELIMINARY;      PRT;   356 AA.
AC   Q918T5;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   E2 protein (Fragment).
GN   E2.
OS   Human papillomavirus type 16.
OC   Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC   Papillomavirus.
OX   NCBI_TaxID=10581;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=HPV16E2CC6;
RA   Watts K.J., Thompson C.H., Cossart Y.E., Rose B.R.;
RT   "Sequence variation and physical state of human Papillomavirus type 16
RT   cervical cancer isolates from Australia and New Caledonia.";
RL   Submitted (AUG-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF407219; AAL01401.1; -.
DR   InterPro; IPR000427; E2_C.
DR   InterPro; IPR001866; E2_N.
DR   Pfam; PF00511; E2_C; 1.
DR   Pfam; PF00508; E2_N; 1.
DR   ProDom; PD000672; E2_C; 1.
DR   ProDom; PD000678; E2_N; 1.
FT   NON_TER       1      1
FT   NON_TER     356    356
SQ   SEQUENCE   356 AA;  40804 MW;  5C0A1B6BE7994EEE CRC64;

  Query Match             95.0%;  Score 113;  DB 12;  Length 356;
  Best Local Similarity   94.7%;  Pred. No. 7.5e-10;
  Matches   18;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      1 DICNTMHYTNWTHIYICEE 19
          |||||||||| |||||||
Db    118 DICNTMHYTNWKHIYICEE 136


RESULT 15
Q918T3
ID   Q918T3      PRELIMINARY;      PRT;   356 AA.
AC   Q918T3;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   E2 protein (Fragment).
GN   E2.
OS   Human papillomavirus type 16.
OC   Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC   Papillomavirus.
OX   NCBI_TaxID=10581;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=HPV16E2CC7;
RA   Watts K.J., Thompson C.H., Cossart Y.E., Rose B.R.;
RT   "Sequence variation and physical state of human Papillomavirus type 16
RT   cervical cancer isolates from Australia and New Caledonia.";
RL   Submitted (AUG-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF407220; AAL01404.1; -.
DR   InterPro; IPR000427; E2_C.
DR   InterPro; IPR001866; E2_N.
DR   Pfam; PF00511; E2_C; 1.
DR   Pfam; PF00508; E2_N; 1.
DR   ProDom; PD000672; E2_C; 1.
DR   ProDom; PD000678; E2_N; 1.
FT   NON_TER       1      1
FT   NON_TER     356    356
SQ   SEQUENCE   356 AA;  40695 MW;  418AAD6FB86A0110 CRC64;

  Query Match             95.0%;  Score 113;  DB 12;  Length 356;
  Best Local Similarity   94.7%;  Pred. No. 7.5e-10;
```

Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy       1 DICNTMHYTNWTHIYICEE 19
           ||||||||||| |||||||
Db     118 DICNTMHYTNWKHIYICEE 136
```

Search completed: January 14, 2003, 18:40:48
Job time : 31 secs

GenCore version 5.1.3

OM protein - protein search, using sw model

Run on: January 14, 2003, 18:38:58 ; Search time 15 Seconds
(without alignments)
121.770 Million cell updates/sec

Title: US-09-828-645-1
Perfect score: 119
Sequence: 1 DICNTMHYTNWTHIYICEE 19

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_73:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	119	100.0	365	1	W2WLHS	E2 protein - human
2	87	73.1	367	1	W2WL35	E2 protein - human
3	87	73.1	367	2	S36524	E2 protein - human
4	79	66.4	372	1	W2WL31	E2 protein - human
5	72	60.5	358	1	W2WL58	E2 protein - human
6	66	55.5	353	1	W2WL33	E2 protein - human
7	64	53.8	377	1	W2WLC1	E2 protein - pygmy
8	62	52.1	398	1	W2WL42	E2 protein - human
9	58.5	49.2	368	2	S36576	E2 protein - human
10	58	48.7	394	2	S36512	E2 protein - human
11	55	46.2	370	2	S36558	E2 protein - human
12	55	46.2	375	2	S36587	E2 protein - human
13	52.5	44.1	377	1	W2WL13	E2 protein - human
14	52	43.7	367	1	W2WL11	E2 protein - human
15	51	42.9	366	1	W2WLR1	E2 protein - rhesu
16	47	39.5	358	1	W2WL51	E2 protein - human
17	47	39.5	368	1	W2WL6	E2 protein - human
18	46	38.7	370	1	W2WL39	E2 protein - human
19	46	38.7	1080	2	T19048	probable Pro-X car
20	45	37.8	154	2	S58070	probable olfactory
21	45	37.8	155	2	T28164	hypothetical prote
22	45	37.8	321	2	S65225	probable membrane
23	45	37.8	345	2	S36518	E2 protein - human
24	45	37.8	382	2	S36476	E2 protein - human
25	45	37.8	383	1	S15624	E2 protein - human
26	45	37.8	388	2	S36500	E2 protein - human
27	44	37.0	99	2	AH3429	transposase BMEI14
28	44	37.0	158	2	T26692	hypothetical prote
29	43	36.1	101	2	AC3467	sarcosine oxidase
30	43	36.1	152	1	MNNZ2G	gene 2 protein - p
31	43	36.1	296	1	A69693	integrase/recombin
32	43	36.1	334	1	W2WLB4	E2 protein - bovin
33	43	36.1	365	2	F88449	protein F54D8.4 [i
34	43	36.1	375	2	S36547	E2 protein - human
35	43	36.1	768	2	B97083	glycosyltransferas
36	43	36.1	929	2	T17392	vrlL protein - Dic
37	43	36.1	1301	1	A41622	protein-tyrosine-p
38	43	36.1	1529	2	T02730	RNA-directed DNA p
39	42.5	35.7	407	2	T08732	hypothetical prote
40	42.5	35.7	461	2	T16161	hypothetical prote
41	42.5	35.7	560	2	T16833	hypothetical prote
42	42.5	35.7	1031	2	T43458	hypothetical prote
43	42	35.3	282	2	JC5677	RNA4 protein - Bee
44	42	35.3	282	2	C44503	p31 protein - beet
45	42	35.3	305	2	T23022	hypothetical prote

ALIGNMENTS

RESULT 1
W2WLHS
E2 protein - human papillomavirus type 16
C;Species: human papillomavirus type 16
C;Date: 28-May-1986 #sequence_revision 28-May-1986 #text_change 20-Aug-1999
C;Accession: A03669; T10429
R;Seedorf, K.; Krammer, G.; Durst, M.; Suhai, S.; Rowekamp, W.G.
Virology 145, 181-185, 1985
A;Title: Human papillomavirus type 16 DNA sequence.
A;Reference number: A22355; MUID:85246220; PMID:2990099
A;Accession: A03669
A;Molecule type: DNA
A;Residues: 1-365 <SEE>
A;Cross-references: GB:K02718; NID:g333031; PIDN:AAA46941.1; PID:g333035
R;Kennedy, I.M.; Haddow, J.K.; Clements, J.B.
J. Virol. 65, 2093-2097, 1991
A;Title: A negative element in the human poapillomavirus type 16 genome acts at the l
A;Reference number: Z17014; MUID:91162763; PMID:1848319
A;Accession: T10429
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-365 <KEN>
A;Cross-references: EMBL:K02718; NID:g333031; PIDN:AAA46941.1; PID:g333035
C;Genetics:
A;Gene: E2
C;Superfamily: papillomavirus E2 protein
C;Keywords: DNA binding; early protein; transcription regulation

Query Match 100.0%; Score 119; DB 1; Length 365;
Best Local Similarity 100.0%; Pred. No. 1.1e-09;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       1 DICNTMHYTNWTHIYICEE 19
           |||||||||||||||||||
Db     124 DICNTMHYTNWTHIYICEE 142
```

RESULT 2
W2WL35
E2 protein - human papillomavirus type 35
C;Species: human papillomavirus type 35
A;Note: host Homo sapiens (man)
C;Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 16-Jul-1999
C;Accession: B40824
R;Marich, J.E.; Pontsler, A.V.; Rice, S.M.; McGraw, K.A.; Dubensky, T.W.
Virology 186, 770-776, 1992
A;Title: The phylogenetic relationship and complete nucleotide sequence of human papi
A;Reference number: A40824; MUID:92124753; PMID:1310198
A;Accession: B40824
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-367 <MAR>

```
A;Cross-references: GB:M74117; NID:g333050; PIDN:AAA46969.1; PID:g333054
C;Superfamily: papillomavirus E2 protein
C;Keywords: DNA binding; early protein; transcription regulation

  Query Match             73.1%;  Score 87;  DB 1;  Length 367;
  Best Local Similarity   87.5%;  Pred. No. 4.4e-05;
  Matches   14;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      4 NTMHYTNWTHIYICEE 19
          |||||||||||| |:
Db    128 NTMHYTNWTHIYILED 143


RESULT 3
S36524
E2 protein - human papillomavirus type 35H
C;Species: human papillomavirus type 35H
C;Date: 09-Dec-1993 #sequence_revision 26-Jul-1996 #text_change 16-Feb-1997
C;Accession: S36524
R;Delius, H.; Hofmann, B.
submitted to the EMBL Data Library, August 1993
A;Description: Primer-directed sequencing of human papillomavirus types.
A;Reference number: S36469
A;Accession: S36524
A;Molecule type: DNA
A;Residues: 1-367 <DEL>
A;Cross-references: EMBL:X74477
A;Experimental source: strain 35H
C;Superfamily: papillomavirus E2 protein
C;Keywords: DNA binding; early protein; transcription regulation

  Query Match             73.1%;  Score 87;  DB 2;  Length 367;
  Best Local Similarity   87.5%;  Pred. No. 4.4e-05;
  Matches   14;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      4 NTMHYTNWTHIYICEE 19
          |||||||||||| |:
Db    128 NTMHYTNWTHIYILED 143


RESULT 4
W2WL31
E2 protein - human papillomavirus type 31
C;Species: human papillomavirus type 31
A;Note: host Homo sapiens (man)
C;Date: 31-Mar-1990 #sequence_revision 31-Mar-1990 #text_change 16-Jul-1999
C;Accession: D32444
R;Goldsborough, M.D.; DiSilvestre, D.; Temple, G.F.; Lorincz, A.T.
Virology 171, 306-311, 1989
A;Title: Nucleotide sequence of human papillomavirus type 31: a cervical neoplasia-assoc
A;Reference number: A94398; MUID:89299478; PMID:2545036
A;Accession: D32444
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-372 <GOL>
A;Cross-references: GB:J04353; NID:g333048; PIDN:AAA46953.1; PID:g459919
C;Superfamily: papillomavirus E2 protein
C;Keywords: DNA binding; early protein; transcription regulation

  Query Match             66.4%;  Score 79;  DB 1;  Length 372;
  Best Local Similarity   70.6%;  Pred. No. 0.00063;
  Matches   12;  Conservative    2;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      1 DICNTMHYTNWTHIYIC 17
          |: ||||||||  ||:|
Db    124 DVHNTMHYTNWKFIYLC 140


RESULT 5
W2WL58
E2 protein - human papillomavirus type 58
C;Species: human papillomavirus type 58
A;Note: host Homo sapiens (man)
C;Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 21-Jul-2000
C;Accession: B36779
R;Kirii, Y.; Iwamoto, S.; Matsukura, T.
Virology 185, 424-427, 1991
A;Title: Human papillomavirus type 58 DNA sequence.
A;Reference number: A36779; MUID:92024102; PMID:1656594
A;Accession: B36779
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-358 <KIR>
A;Cross-references: GB:D90400; NID:g222386; PIDN:BAA31848.1; PID:g3337101
C;Superfamily: papillomavirus E2 protein
C;Keywords: DNA binding; early protein; transcription regulation

  Query Match             60.5%;  Score 72;  DB 1;  Length 358;
  Best Local Similarity   68.4%;  Pred. No. 0.0061;
  Matches   13;  Conservative    1;  Mismatches    5;  Indels    0;  Gaps    0;

Qy      1 DICNTMHYTNWTHIYICEE 19
          |  ||| ||||: ||| ||
Db    124 DKANTMDYTNWSEIYIIEE 142


RESULT 6
W2WL33
E2 protein - human papillomavirus type 33
C;Species: human papillomavirus type 33
C;Date: 30-Jun-1987 #sequence_revision 30-Jun-1987 #text_change 16-Jul-1999
C;Accession: A03670
R;Cole, S.T.; Streeck, R.E.
J. Virol. 58, 991-995, 1986
A;Title: Genome organization and nucleotide sequence of human papillomavirus type 33,
A;Reference number: A93020; MUID:86200464; PMID:3009902
A;Accession: A03670
A;Molecule type: DNA
A;Residues: 1-353 <COL>
A;Cross-references: GB:M12732; NID:g333049; PIDN:AAA46961.1; PID:g463180
C;Superfamily: papillomavirus E2 protein
C;Keywords: DNA binding; early protein; transcription regulation

  Query Match             55.5%;  Score 66;  DB 1;  Length 353;
  Best Local Similarity   75.0%;  Pred. No. 0.044;
  Matches   12;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      4 NTMHYTNWTHIYICEE 19
          ||| ||||  ||| ||
Db    127 NTMDYTNWGEIYIIEE 142


RESULT 7
W2WLC1
E2 protein - pygmy chimpanzee papillomavirus (type 1)
C;Species: pygmy chimpanzee papillomavirus
C;Date: 30-Jun-1993 #sequence_revision 30-Jun-1993 #text_change 16-Jul-1999
C;Accession: D36818
R;van Ranst, M.; Fuse, A.; Fiten, P.; Beuken, E.; Pfister, H.; Burk, R.D.; Opdenakker
Virology 190, 587-596, 1992
A;Title: Human papillomavirus type 13 and pygmy chimpanzee papillomavirus type 1: Com
A;Reference number: A42955; MUID:92391075; PMID:1325697
A;Accession: D36818
A;Molecule type: DNA
A;Residues: 1-377 <VAN>
A;Cross-references: EMBL:X62844; NID:g61010; PIDN:CAA44658.1; PID:g61014
C;Superfamily: papillomavirus E2 protein
C;Keywords: DNA binding; early protein; transcription regulation

  Query Match             53.8%;  Score 64;  DB 1;  Length 377;
  Best Local Similarity   60.0%;  Pred. No. 0.091;
  Matches    9;  Conservative    3;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      4 NTMHYTNWTHIYICE 18
          |:|||  | :||:||
Db    127 NSMHYVLWKYIYVCE 141
```

```
RESULT 8
W2WL42
E2 protein - human papillomavirus type 42
C;Species: human papillomavirus type 42
A;Note: host Homo sapiens (man)
C;Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 16-Feb-1997
C;Accession: B39451
R;Philipp, W.; Honore, N.; Sapp, M.; Cole, S.T.; Streeck, R.E.
Virology 186, 331-334, 1992
A;Title: Human papillomavirus type 42: new sequence, conserved genome organization.
A;Reference number: A39451; MUID:92087479; PMID:1309278
A;Accession: B39451
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-398 <PHI>
A;Cross-references: GB:M73236
C;Superfamily: papillomavirus E2 protein
C;Keywords: DNA binding; early protein; transcription regulation

  Query Match             52.1%;  Score 62;  DB 1;  Length 398;
  Best Local Similarity   76.9%;  Pred. No. 0.19;
  Matches   10;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      4 NTMHYTNWTHIYI 16
          | |||| ||:|||
Db    127 NAMHYTAWTYIYI 139


RESULT 9
S36576
E2 protein - human papillomavirus type 52
C;Species: human papillomavirus type 52
C;Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 26-Aug-1999
C;Accession: S36576
R;Delius, H.; Hofmann, B.
submitted to the EMBL Data Library, August 1993
A;Description: Primer-directed sequencing of human papillomavirus types.
A;Reference number: S36469
A;Accession: S36576
A;Molecule type: DNA
A;Residues: 1-368 <DEL>
A;Cross-references: EMBL:X74481; NID:g397038; PIDN:CAA52588.1; PID:g397042
C;Superfamily: papillomavirus E2 protein
C;Keywords: DNA binding; early protein; transcription regulation

  Query Match             49.2%;  Score 58.5;  DB 2;  Length 368;
  Best Local Similarity   61.1%;  Pred. No. 0.55;
  Matches   11;  Conservative    1;  Mismatches    3;  Indels    3;  Gaps    1;

Qy      4 NTMHYTNWTHIYI---CE 18
          ||| ||||  ||:   ||
Db    127 NTMDYTNWKEIYLLGECE 144


RESULT 10
S36512
E2 protein - human papillomavirus type 32
C;Species: human papillomavirus type 32
C;Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 26-Aug-1999
C;Accession: S36512
R;Delius, H.; Hofmann, B.
submitted to the EMBL Data Library, August 1993
A;Description: Primer-directed sequencing of human papillomavirus types.
A;Reference number: S36469
A;Accession: S36512
A;Molecule type: DNA
A;Residues: 1-394 <DEL>
A;Cross-references: EMBL:X74475; NID:g396981; PIDN:CAA52552.1; PID:g396985
C;Superfamily: papillomavirus E2 protein
C;Keywords: DNA binding; early protein; transcription regulation

  Query Match             48.7%;  Score 58;  DB 2;  Length 394;
  Best Local Similarity   69.2%;  Pred. No. 0.69;
  Matches    9;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      4 NTMHYTNWTHIYI 16
          | |||| || ||:
Db    127 NAMHYTAWTFIYV 139


RESULT 11
S36558
E2 protein - human papillomavirus type 40
C;Species: human papillomavirus type 40
C;Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 26-Aug-1999
C;Accession: S36558
R;Delius, H.; Hofmann, B.
submitted to the EMBL Data Library, August 1993
A;Description: Primer-directed sequencing of human papillomavirus types.
A;Reference number: S36469
A;Accession: S36558
A;Molecule type: DNA
A;Residues: 1-370 <DEL>
A;Cross-references: EMBL:X74478; NID:g397014; PIDN:CAA52570.1; PID:g397018
C;Superfamily: papillomavirus E2 protein
C;Keywords: DNA binding; early protein; transcription regulation

  Query Match             46.2%;  Score 55;  DB 2;  Length 370;
  Best Local Similarity   61.5%;  Pred. No. 1.8;
  Matches    8;  Conservative    2;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      4 NTMHYTNWTHIYI 16
          | |||| || :|:
Db    127 NAMHYTLWTTVYV 139


RESULT 12
S36587
E2 protein - human papillomavirus type 7
C;Species: human papillomavirus type 7
C;Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 26-Aug-1999
C;Accession: S36587
R;Delius, H.; Hofmann, B.
submitted to the EMBL Data Library, August 1993
A;Description: Primer-directed sequencing of human papillomavirus types.
A;Reference number: S36469
A;Accession: S36587
A;Molecule type: DNA
A;Residues: 1-375 <DEL>
A;Cross-references: EMBL:X74463; NID:g397060; PIDN:CAA52479.1; PID:g397064
C;Superfamily: papillomavirus E2 protein
C;Keywords: DNA binding; early protein; transcription regulation

  Query Match             46.2%;  Score 55;  DB 2;  Length 375;
  Best Local Similarity   61.5%;  Pred. No. 1.8;
  Matches    8;  Conservative    2;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      4 NTMHYTNWTHIYI 16
          | |||| || :|:
Db    127 NAMHYTLWTAVYV 139


RESULT 13
W2WL13
E2 protein - human papillomavirus type 13
C;Species: human papillomavirus type 13
A;Note: host Homo sapiens (man)
C;Date: 30-Jun-1993 #sequence_revision 30-Jun-1993 #text_change 16-Jul-1999
C;Accession: D42955
R;van Ranst, M.; Fuse, A.; Fiten, P.; Beuken, E.; Pfister, H.; Burk, R.D.; Opdenakker
Virology 190, 587-596, 1992
A;Title: Human papillomavirus type 13 and pygmy chimpanzee papillomavirus type 1: Com
A;Reference number: A42955; MUID:92391075; PMID:1325697
A;Accession: D42955
```

```
A;Molecule type: DNA
A;Residues: 1-377 <VAN>
A;Cross-references: EMBL:X62843; NID:g60295; PIDN:CAA44650.1; PID:g60299
C;Superfamily: papillomavirus E2 protein
C;Keywords: DNA binding; early protein; transcription regulation

  Query Match             44.1%;  Score 52.5;  DB 1;  Length 377;
  Best Local Similarity   52.9%;  Pred. No. 4.1;
  Matches    9;  Conservative    3;  Mismatches    2;  Indels    3;  Gaps    1;

Qy       3 CNT---MHYTNWTHIYI 16
           |||   | | :||:||:
Db     123 CNTDNRMDYVSWTYIYV 139


RESULT 14
W2WL11
E2 protein - human papillomavirus type 11
C;Species: human papillomavirus type 11
C;Date: 13-Aug-1986 #sequence_revision 13-Aug-1986 #text_change 16-Jul-1999
C;Accession: A03668
R;Dartmann, K.; Schwarz, E.; Gissmann, L.; zur Hausen, H.
Virology 151, 124-130, 1986
A;Title: The nucleotide sequence and genome organization of human papilloma virus type 1
A;Reference number: A94338; MUID:86181601; PMID:3008427
A;Accession: A03668
A;Molecule type: DNA
A;Residues: 1-367 <DAR>
A;Cross-references: GB:M14119; NID:g333026; PIDN:AAA46930.1; PID:g496196
C;Superfamily: papillomavirus E2 protein
C;Keywords: DNA binding; early protein; transcription regulation

  Query Match             43.7%;  Score 52;  DB 1;  Length 367;
  Best Local Similarity   61.5%;  Pred. No. 4.7;
  Matches    8;  Conservative    1;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       4 NTMHYTNWTHIYI 16
           | | |  |||||:
Db     127 NVMEYVVWTHIYL 139


RESULT 15
W2WLR1
E2 protein - rhesus papillomavirus (type 1)
C;Species: rhesus papillomavirus
C;Date: 31-Dec-1991 #sequence_revision 31-Dec-1991 #text_change 16-Feb-1997
C;Accession: D38503
R;Ostrow, R.S.; LaBresh, K.V.; Faras, A.J.
Virology 181, 424-429, 1991
A;Title: Characterization of the complete RhPV 1 genomic sequence and an integration loc
A;Reference number: A38503; MUID:91135018; PMID:1847267
A;Accession: D38503
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-366 <OST>
A;Cross-references: EMBL:M37717
C;Superfamily: papillomavirus E2 protein
C;Keywords: DNA binding; early protein; transcription regulation

  Query Match             42.9%;  Score 51;  DB 1;  Length 366;
  Best Local Similarity   61.5%;  Pred. No. 6.5;
  Matches    8;  Conservative    1;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       4 NTMHYTNWTHIYI 16
           ||| |  | |||:
Db     128 NTMEYVLWGHIYV 140


Search completed: January 14, 2003, 18:41:10
Job time : 16 secs
```

GenCore version 5.1.3

OM protein - protein search, using sw model

Run on: January 14, 2003, 18:38:03 ; Search time 11 Seconds
(without alignments)
71.641 Million cell updates/sec

Title: US-09-828-645-1
Perfect score: 119
Sequence: 1 DICNTMHYTNWTHIYICEE 19

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	119	100.0	365	1	VE2_HPV16	P03120 human papil
2	87	73.1	367	1	VE2_HPV35	P27222 human papil
3	79	66.4	372	1	VE2_HPV31	P17383 human papil
4	72	60.5	358	1	VE2_HPV58	P26546 human papil
5	66	55.5	353	1	VE2_HPV33	P06423 human papil
6	64	53.8	377	1	VE2_PCPV1	Q02264 pygmy chimp
7	62	52.1	398	1	VE2_HPV42	P27223 human papil
8	58.5	49.2	368	1	VE2_HPV52	P36796 human papil
9	58	48.7	394	1	VE2_HPV32	P36791 human papil
10	55	46.2	370	1	VE2_HPV40	P36793 human papil
11	55	46.2	375	1	VE2_HPV07	P36779 human papil
12	52.5	44.1	377	1	VE2_HPV13	Q02263 human papil
13	52	43.7	367	1	VE2_HPV11	P04015 human papil
14	51	42.9	366	1	VE2_RHPV1	P22156 rhesus papi
15	50	42.0	404	1	VE2_HPV60	Q80944 human papil
16	48	40.3	402	1	VE2_HPV65	Q07851 human papil
17	47	39.5	358	1	VE2_HPV51	P26547 human papil
18	47	39.5	368	1	VE2_HPV6A	Q84294 human papil
19	47	39.5	368	1	VE2_HPV6B	P03119 human papil
20	47	39.5	441	1	VE2_HPV38	Q80910 human papil
21	46	38.7	370	1	VE2_HPV39	P24830 human papil
22	46	38.7	509	1	VE2_HPV36	P50809 human papil
23	45	37.8	345	1	VE2_HPV34	P36792 human papil
24	45	37.8	382	1	VE2_HPV15	P36784 human papil
25	45	37.8	383	1	VE2_HPV57	P22155 human papil
26	45	37.8	388	1	VE2_HPV27	P36789 human papil
27	44	37.0	378	1	VE2_HPV55	Q80937 human papil
28	44	37.0	402	1	VE2_HPV04	Q07849 human papil
29	44	37.0	436	1	VE2_HPV22	P50768 human papil
30	44	37.0	663	1	SCG2_XENLA	O13263 xenopus lae
31	43	36.1	152	1	VNS2_PVM	P28889 pneumonia v
32	43	36.1	296	1	RIPX_BACSU	P46352 bacillus su
33	43	36.1	334	1	VE2_BPV4	P08345 bovine papi
34	43	36.1	375	1	VE2_HPV26	P36788 human papil
35	43	36.1	431	1	VE2_HPV23	P50769 human papil
36	43	36.1	1301	1	PTP9_DROME	P35832 drosophila
37	42.5	35.7	461	1	YUL2_CAEEL	Q19802 caenorhabdi
38	42.5	35.7	1066	1	Z295_HUMAN	Q9ulj3 homo sapien
39	42	35.3	282	1	Y32K_BNYVG	P19231 beet necrot
40	42	35.3	360	1	VE2_HPV70	P50773 human papil
41	42	35.3	416	1	RHLC_GORGO	Q28426 gorilla gor
42	42	35.3	499	1	TMOA_PSEME	Q00456 pseudomonas
43	41.5	34.9	2731	1	RRPB_CVMJH	P29982 murine coro
44	41	34.5	199	1	YE65_MYCPN	P75318 mycoplasma
45	41	34.5	225	1	UNG_BACSU	P39615 bacillus su

ALIGNMENTS

RESULT 1
VE2_HPV16
ID VE2_HPV16 STANDARD; PRT; 365 AA.
AC P03120;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE Regulatory protein E2.
GN E2.
OS Human papillomavirus type 16.
OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC Papillomavirus.
OX NCBI_TaxID=10581;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=85246220; PubMed=2990099;
RA Seedorf K., Krammer G., Durst M., Suhai S., Rowekamp W.G.;
RT "Human papillomavirus type 16 DNA sequence.";
RL Virology 145:181-185(1985).
RN [2]
RP FUNCTION.
RX MEDLINE=87198893; PubMed=3033289;
RA Phelps W.C., Howley P.M.;
RT "Transcriptional trans-activation by the human papillomavirus type 16
RT E2 gene product.";
RL J. Virol. 61:1630-1638(1987).
CC -!- FUNCTION: E2 REGULATES VIRAL TRANSCRIPTION AND DNA REPLICATION.
CC IT BINDS TO THE E2RE RESPONSE ELEMENT (5'-ACCNNNNNNGGT-3') PRESENT
CC IN MULTIPLE COPIES IN THE REGULATORY REGION. IT CAN EITHER
CC ACTIVATE OR REPRESS TRANSCRIPTION DEPENDING OF E2RE'S POSITION
CC WITH REGARDS TO PROXIMAL PROMOTER ELEMENTS. REPRESSION OCCURS
CC BY STERICALLY HINDERING THE ASSEMBLY OF THE TRANSCRIPTION
CC INITIATION COMPLEX. THE E1-E2 COMPLEX BINDS TO THE ORIGIN OF DNA
CC REPLICATION.
CC -!- SUBUNIT: BINDS DNA AS A DIMER.
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- DISEASE: HPV16, IN COMPARISON TO HPV TYPES 6 AND 11, IS MORE
CC OFTEN ASSOCIATED WITH MALIGNANT GENITAL CANCERS IN HUMANS.
CC ----------

CC ----------
DR EMBL; K02718; AAA46941.1; -.
DR PIR; A03669; W2WLHS.
DR HSSP; P17383; 1DHM.
DR InterPro; IPR000427; E2_C.
DR InterPro; IPR001866; E2_N.
DR Pfam; PF00508; E2_N; 1.
DR Pfam; PF00511; E2_C; 1.
DR ProDom; PD000672; E2_C; 1.
DR ProDom; PD000678; E2_N; 1.

```
KW   Early protein; Transcription regulation; Activator; DNA-binding;
KW   Trans-acting factor; DNA replication; Repressor; Nuclear protein.
SQ   SEQUENCE   365 AA;  41825 MW;  27581F82B246E112 CRC64;

  Query Match             100.0%;  Score 119;  DB 1;  Length 365;
  Best Local Similarity   100.0%;  Pred. No. 9.2e-10;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 DICNTMHYTNWTHIYICEE 19
          ||||||||||||||||||||
Db    124 DICNTMHYTNWTHIYICEE 142


RESULT 2
VE2_HPV35
ID   VE2_HPV35      STANDARD;      PRT;   367 AA.
AC   P27222;
DT   01-AUG-1992 (Rel. 23, Created)
DT   01-JUN-1994 (Rel. 29, Last sequence update)
DT   15-JUL-1998 (Rel. 36, Last annotation update)
DE   Regulatory protein E2.
GN   E2.
OS   Human papillomavirus type 35.
OC   Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC   Papillomavirus.
OX   NCBI_TaxID=10587;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=94265501; PubMed=8205838;
RA   Delius H., Hofmann B.;
RT   "Primer-directed sequencing of human papillomavirus types.";
RL   Curr. Top. Microbiol. Immunol. 186:13-31(1994).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=92124753; PubMed=1310198;
RA   Marich J.E., Pontsler A.V., Rice S.M.,. McGraw K.A., Dubensky T.W.;
RT   "The phylogenetic relationship and complete nucleotide sequence of
RT   human papillomavirus type 35.";
RL   Virology 186:770-776(1992).
CC   -!- FUNCTION: E2 REGULATES VIRAL TRANSCRIPTION AND DNA REPLICATION.
CC       IT BINDS TO THE E2RE RESPONSE ELEMENT (5'-ACCNNNNNNGGT-3') PRESENT
CC       IN MULTIPLE COPIES IN THE REGULATORY REGION. IT CAN EITHER
CC       ACTIVATE OR REPRESS TRANSCRIPTION DEPENDING OF E2RE'S POSITION
CC       WITH REGARDS TO PROXIMAL PROMOTER ELEMENTS. REPRESSION OCCURS
CC       BY STERICALLY HINDERING THE ASSEMBLY OF THE TRANSCRIPTION
CC       INITIATION COMPLEX. THE E1-E2 COMPLEX BINDS TO THE ORIGIN OF DNA
CC       REPLICATION.
CC   -!- SUBUNIT: BINDS DNA AS A DIMER.
CC   -!- SUBCELLULAR LOCATION: Nuclear.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X74477; CAA52564.1; ALT_TERM.
DR   EMBL; M74117; AAA46969.1; -.
DR   PIR; B40824; W2WL35.
DR   PIR; S36524; S36524.
DR   HSSP; P17383; 1DHM.
DR   InterPro; IPR000427; E2_C.
DR   InterPro; IPR001866; E2_N.
DR   Pfam; PF00508; E2_N; 1.
DR   Pfam; PF00511; E2_C; 1.
DR   ProDom; PD000672; E2_C; 1.
DR   ProDom; PD000678; E2_N; 1.
KW   Early protein; Transcription regulation; Activator; DNA-binding;
KW   Trans-acting factor; DNA replication; Repressor; Nuclear protein.
FT   CONFLICT     92     92       T -> D (IN REF. 2).
FT   CONFLICT    108    109       QG -> TR (IN REF. 2).
FT   CONFLICT    111    111       F -> L (IN REF. 2).
FT   CONFLICT    114    116       HGV -> DVY (IN REF. 2).
FT   CONFLICT    120    120       V -> A (IN REF. 2).
FT   CONFLICT    303    303       L -> S (IN REF. 2).
SQ   SEQUENCE   367 AA;  41897 MW;  747C197E02C0AF35 CRC64;

  Query Match             73.1%;  Score 87;  DB 1;  Length 367;
  Best Local Similarity   87.5%;  Pred. No. 2.9e-05;
  Matches   14;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      4 NTMHYTNWTHIYICEE 19
          |||||||||||||  |:
Db    128 NTMHYTNWTHIYILED 143


RESULT 3
VE2_HPV31
ID   VE2_HPV31      STANDARD;      PRT;   372 AA.
AC   P17383;
DT   01-AUG-1990 (Rel. 15, Created)
DT   01-AUG-1990 (Rel. 15, Last sequence update)
DT   15-JUL-1998 (Rel. 36, Last annotation update)
DE   Regulatory protein E2.
GN   E2.
OS   Human papillomavirus type 31.
OC   Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC   Papillomavirus.
OX   NCBI_TaxID=10585;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=89299478; PubMed=2545036;
RA   Goldsborough M.D., Disilvestre D., Temple G.F., Lorincz A.T.;
RT   "Nucleotide sequence of human papillomavirus type 31: a cervical
RT   neoplasia-associated virus.";
RL   Virology 171:306-311(1989).
RN   [2]
RP   STRUCTURE BY NMR OF 291-372.
RX   MEDLINE=96194130; PubMed=8652551;
RA   Liang H., Petros A.M., Meadows R.P., Yoon H.S., Egan D.A., Walter K.,
RA   Holzman T.F., Robins T., Fesik S.W.;
RT   "Solution structure of the DNA-binding domain of a human
RT   papillomavirus E2 protein: evidence for flexible DNA-binding
RT   regions.";
RL   Biochemistry 35:2095-2103(1996).
CC   -!- FUNCTION: E2 REGULATES VIRAL TRANSCRIPTION AND DNA REPLICATION.
CC       IT BINDS TO THE E2RE RESPONSE ELEMENT (5'-ACCNNNNNNGGT-3') PRESENT
CC       IN MULTIPLE COPIES IN THE REGULATORY REGION. IT CAN EITHER
CC       ACTIVATE OR REPRESS TRANSCRIPTION DEPENDING OF E2RE'S POSITION
CC       WITH REGARDS TO PROXIMAL PROMOTER ELEMENTS. REPRESSION OCCURS
CC       BY STERICALLY HINDERING THE ASSEMBLY OF THE TRANSCRIPTION
CC       INITIATION COMPLEX. THE E1-E2 COMPLEX BINDS TO THE ORIGIN OF DNA
CC       REPLICATION.
CC   -!- SUBUNIT: BINDS DNA AS A DIMER.
CC   -!- SUBCELLULAR LOCATION: Nuclear.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; J04353; AAA46953.1; -.
DR   PIR; D32444; W2WL31.
DR   PDB; 1DHM; 07-DEC-96.
DR   InterPro; IPR000427; E2_C.
DR   InterPro; IPR001866; E2_N.
DR   Pfam; PF00508; E2_N; 1.
DR   Pfam; PF00511; E2_C; 1.
DR   ProDom; PD000672; E2_C; 1.
DR   ProDom; PD000678; E2_N; 1.
KW   Early protein; Transcription regulation; Activator; DNA-binding;
```

```
KW   Trans-acting factor; DNA replication; Repressor; Nuclear protein;
KW   3D-structure.
SQ   SEQUENCE   372 AA;  42104 MW;  4F387C5E3AFB9D1D CRC64;

  Query Match           66.4%;  Score 79;  DB 1;  Length 372;
  Best Local Similarity  70.6%;  Pred. No. 0.00038;
  Matches   12;  Conservative    2;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      1 DICNTMHYTNWTHIYIC 17
          |: |||||||  ||:|
Db    124 DVHNTMHYTNWKFIYLC 140


RESULT 4
VE2_HPV58
ID   VE2_HPV58      STANDARD;      PRT;   358 AA.
AC   P26546;
DT   01-AUG-1992 (Rel. 23, Created)
DT   01-AUG-1992 (Rel. 23, Last sequence update)
DT   15-JUL-1998 (Rel. 36, Last annotation update)
DE   Regulatory protein E2.
GN   E2.
OS   Human papillomavirus type 58.
OC   Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC   Papillomavirus.
OX   NCBI_TaxID=10598;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=92024102; PubMed=1656594;
RA   Kirii Y., Iwamoto S., Matsukura T.;
RT   "Human papillomavirus type 58 DNA sequence.";
RL   Virology 185:424-427(1991).
CC   -!- FUNCTION: E2 REGULATES VIRAL TRANSCRIPTION AND DNA REPLICATION.
CC       IT BINDS TO THE E2RE RESPONSE ELEMENT (5'-ACCNNNNNNGGT-3') PRESENT
CC       IN MULTIPLE COPIES IN THE REGULATORY REGION. IT CAN EITHER
CC       ACTIVATE OR REPRESS TRANSCRIPTION DEPENDING OF E2RE'S POSITION
CC       WITH REGARDS TO PROXIMAL PROMOTER ELEMENTS. REPRESSION OCCURS
CC       BY STERICALLY HINDERING THE ASSEMBLY OF THE TRANSCRIPTION
CC       INITIATION COMPLEX. THE E1-E2 COMPLEX BINDS TO THE ORIGIN OF DNA
CC       REPLICATION.
CC   -!- SUBUNIT: BINDS DNA AS A DIMER.
CC   -!- SUBCELLULAR LOCATION: Nuclear.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; D90400; BAA31848.1; -.
DR   PIR; B36779; W2WL58.
DR   HSSP; P17383; 1DHM.
DR   InterPro; IPR000427; E2_C.
DR   InterPro; IPR001866; E2_N.
DR   Pfam; PF00508; E2_N; 1.
DR   Pfam; PF00511; E2_C; 1.
DR   ProDom; PD000672; E2_C; 1.
DR   ProDom; PD000678; E2_N; 1.
KW   Early protein; Transcription regulation; Activator; DNA-binding;
KW   Trans-acting factor; DNA replication; Repressor; Nuclear protein.
SQ   SEQUENCE   358 AA;  40781 MW;  A3CA0BE001E2BD1E CRC64;

  Query Match           60.5%;  Score 72;  DB 1;  Length 358;
  Best Local Similarity  68.4%;  Pred. No. 0.0035;
  Matches   13;  Conservative    1;  Mismatches    5;  Indels    0;  Gaps    0;

Qy      1 DICNTMHYTNWTHIYICEE 19
          |  ||| ||||: ||| ||
Db    124 DKANTMDYTNWSEIYIIEE 142


RESULT 5
VE2_HPV33
ID   VE2_HPV33      STANDARD;      PRT;   353 AA.
AC   P06423;
DT   01-JAN-1988 (Rel. 06, Created)
DT   01-JAN-1988 (Rel. 06, Last sequence update)
DT   15-JUL-1998 (Rel. 36, Last annotation update)
DE   Regulatory protein E2.
GN   E2.
OS   Human papillomavirus type 33.
OC   Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC   Papillomavirus.
OX   NCBI_TaxID=10586;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=86200464; PubMed=3009902;
RA   Cole S.T., Streeck R.E.;
RT   "Genome organization and nucleotide sequence of human papillomavirus
RT   type 33, which is associated with cervical cancer.";
RL   J. Virol. 58:991-995(1986).
CC   -!- FUNCTION: E2 REGULATES VIRAL TRANSCRIPTION AND DNA REPLICATION.
CC       IT BINDS TO THE E2RE RESPONSE ELEMENT (5'-ACCNNNNNNGGT-3') PRESENT
CC       IN MULTIPLE COPIES IN THE REGULATORY REGION. IT CAN EITHER
CC       ACTIVATE OR REPRESS TRANSCRIPTION DEPENDING OF E2RE'S POSITION
CC       WITH REGARDS TO PROXIMAL PROMOTER ELEMENTS. REPRESSION OCCURS
CC       BY STERICALLY HINDERING THE ASSEMBLY OF THE TRANSCRIPTION
CC       INITIATION COMPLEX. THE E1-E2 COMPLEX BINDS TO THE ORIGIN OF DNA
CC       REPLICATION.
CC   -!- SUBUNIT: BINDS DNA AS A DIMER.
CC   -!- SUBCELLULAR LOCATION: Nuclear.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M12732; AAA46961.1; -.
DR   PIR; A03670; W2WL33.
DR   HSSP; P17383; 1DHM.
DR   InterPro; IPR000427; E2_C.
DR   InterPro; IPR001866; E2_N.
DR   Pfam; PF00508; E2_N; 1.
DR   Pfam; PF00511; E2_C; 1.
DR   ProDom; PD000672; E2_C; 1.
DR   ProDom; PD000678; E2_N; 1.
KW   Early protein; Transcription regulation; Activator; DNA-binding;
KW   Trans-acting factor; DNA replication; Repressor; Nuclear protein.
SQ   SEQUENCE   353 AA;  40253 MW;  673A9D765DBB11BC CRC64;

  Query Match           55.5%;  Score 66;  DB 1;  Length 353;
  Best Local Similarity  75.0%;  Pred. No. 0.024;
  Matches   12;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      4 NTMHYTNWTHIYICEE 19
          ||| ||||  ||| ||
Db    127 NTMDYTNWGEIYIIEE 142


RESULT 6
VE2_PCPV1
ID   VE2_PCPV1      STANDARD;      PRT;   377 AA.
AC   Q02264;
DT   01-APR-1993 (Rel. 25, Created)
DT   01-APR-1993 (Rel. 25, Last sequence update)
DT   15-JUL-1998 (Rel. 36, Last annotation update)
DE   Regulatory protein E2.
GN   E2.
OS   Pygmy chimpanzee papillomavirus type 1.
OC   Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC   Papillomavirus.
```

```
OX   NCBI_TaxID=10576;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=92391075; PubMed=1325697;
RA   van Ranst M., Fuse A., Fiten P., Beuken E., Pfister H., Burk R.D.,
RA   Opdenakker G.;
RT   "Human papillomavirus type 13 and pygmy chimpanzee papillomavirus
RT   type 1: comparison of the genome organizations.";
RL   Virology 190:587-596(1992).
CC   -!- FUNCTION: E2 REGULATES VIRAL TRANSCRIPTION AND DNA REPLICATION.
CC       IT BINDS TO THE E2RE RESPONSE ELEMENT (5'-ACCNNNNNNGGT-3') PRESENT
CC       IN MULTIPLE COPIES IN THE REGULATORY REGION. IT CAN EITHER
CC       ACTIVATE OR REPRESS TRANSCRIPTION DEPENDING OF E2RE'S POSITION
CC       WITH REGARDS TO PROXIMAL PROMOTER ELEMENTS. REPRESSION OCCURS
CC       BY STERICALLY HINDERING THE ASSEMBLY OF THE TRANSCRIPTION
CC       INITIATION COMPLEX. THE E1-E2 COMPLEX BINDS TO THE ORIGIN OF DNA
CC       REPLICATION.
CC   -!- SUBUNIT: BINDS DNA AS A DIMER.
CC   -!- SUBCELLULAR LOCATION: Nuclear.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X62844; CAA44658.1; -.
DR   PIR; D36818; W2WLC1.
DR   HSSP; P17383; 1DHM.
DR   InterPro; IPR000427; E2_C.
DR   InterPro; IPR001866; E2_N.
DR   Pfam; PF00508; E2_N; 1.
DR   Pfam; PF00511; E2_C; 1.
DR   ProDom; PD000672; E2_C; 1.
DR   ProDom; PD000678; E2_N; 1.
KW   Early protein; Transcription regulation; Activator; DNA-binding;
KW   Trans-acting factor; DNA replication; Repressor; Nuclear protein.
SQ   SEQUENCE   377 AA;  43264 MW;  B04A6A0DB47FF2CA CRC64;

  Query Match             53.8%;  Score 64;  DB 1;  Length 377;
  Best Local Similarity   60.0%;  Pred. No. 0.049;
  Matches    9;  Conservative    3;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       4 NTMHYTNWTHIYICE 18
           |:|||  | :||:||
Db     127 NSMHYVLWKYIYVCE 141


RESULT 7
VE2_HPV42
ID   VE2_HPV42      STANDARD;      PRT;   398 AA.
AC   P27223;
DT   01-AUG-1992 (Rel. 23, Created)
DT   01-AUG-1992 (Rel. 23, Last sequence update)
DT   15-JUL-1998 (Rel. 36, Last annotation update)
DE   Regulatory protein E2.
GN   E2.
OS   Human papillomavirus type 42.
OC   Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC   Papillomavirus.
OX   NCBI_TaxID=10590;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=92087479; PubMed=1309278;
RA   Philipp W., Honore N., Sapp M., Cole S.T., Streeck R.E.;
RT   "Human papillomavirus type 42: new sequences, conserved genome
RT   organization.";
RL   Virology 186:331-334(1992).
CC   -!- FUNCTION: E2 REGULATES VIRAL TRANSCRIPTION AND DNA REPLICATION.
CC       IT BINDS TO THE E2RE RESPONSE ELEMENT (5'-ACCNNNNNNGGT-3') PRESENT
CC       IN MULTIPLE COPIES IN THE REGULATORY REGION. IT CAN EITHER
CC       ACTIVATE OR REPRESS TRANSCRIPTION DEPENDING OF E2RE'S POSITION
CC       WITH REGARDS TO PROXIMAL PROMOTER ELEMENTS. REPRESSION OCCURS
CC       BY STERICALLY HINDERING THE ASSEMBLY OF THE TRANSCRIPTION
CC       INITIATION COMPLEX. THE E1-E2 COMPLEX BINDS TO THE ORIGIN OF DNA
CC       REPLICATION.
CC   -!- SUBUNIT: BINDS DNA AS A DIMER.
CC   -!- SUBCELLULAR LOCATION: Nuclear.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M73236; AAA47044.1; ALT_INIT.
DR   PIR; B39451; W2WL42.
DR   HSSP; P17383; 1DHM.
DR   InterPro; IPR000427; E2_C.
DR   InterPro; IPR001866; E2_N.
DR   Pfam; PF00508; E2_N; 1.
DR   Pfam; PF00511; E2_C; 1.
DR   ProDom; PD000672; E2_C; 1.
DR   ProDom; PD000678; E2_N; 1.
KW   Early protein; Transcription regulation; Activator; DNA-binding;
KW   Trans-acting factor; DNA replication; Repressor; Nuclear protein.
SQ   SEQUENCE   398 AA;  45309 MW;  4D41D7196372808C CRC64;

  Query Match             52.1%;  Score 62;  DB 1;  Length 398;
  Best Local Similarity   76.9%;  Pred. No. 0.099;
  Matches   10;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       4 NTMHYTNWTHIYI 16
           | |||| ||:|||
Db     127 NAMHYTAWTYIYI 139


RESULT 8
VE2_HPV52
ID   VE2_HPV52      STANDARD;      PRT;   368 AA.
AC   P36796;
DT   01-JUN-1994 (Rel. 29, Created)
DT   01-JUN-1994 (Rel. 29, Last sequence update)
DT   15-JUL-1998 (Rel. 36, Last annotation update)
DE   Regulatory protein E2.
GN   E2.
OS   Human papillomavirus type 52.
OC   Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC   Papillomavirus.
OX   NCBI_TaxID=10618;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=94265501; PubMed=8205838;
RA   Delius H., Hofmann B.;
RT   "Primer-directed sequencing of human papillomavirus types.";
RL   Curr. Top. Microbiol. Immunol. 186:13-31(1994).
CC   -!- FUNCTION: E2 REGULATES VIRAL TRANSCRIPTION AND DNA REPLICATION.
CC       IT BINDS TO THE E2RE RESPONSE ELEMENT (5'-ACCNNNNNNGGT-3') PRESENT
CC       IN MULTIPLE COPIES IN THE REGULATORY REGION. IT CAN EITHER
CC       ACTIVATE OR REPRESS TRANSCRIPTION DEPENDING OF E2RE'S POSITION
CC       WITH REGARDS TO PROXIMAL PROMOTER ELEMENTS. REPRESSION OCCURS
CC       BY STERICALLY HINDERING THE ASSEMBLY OF THE TRANSCRIPTION
CC       INITIATION COMPLEX. THE E1-E2 COMPLEX BINDS TO THE ORIGIN OF DNA
CC       REPLICATION.
CC   -!- SUBUNIT: BINDS DNA AS A DIMER.
CC   -!- SUBCELLULAR LOCATION: Nuclear.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
```

```
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; X74481; CAA52588.1; -.
DR   PIR; S36576; S36576.
DR   HSSP; P17383; 1DHM.
DR   InterPro; IPR000427; E2_C.
DR   InterPro; IPR001866; E2_N.
DR   Pfam; PF00508; E2_N; 1.
DR   Pfam; PF00511; E2_C; 1.
DR   ProDom; PD000672; E2_C; 1.
DR   ProDom; PD000678; E2_N; 1.
KW   Early protein; Transcription regulation; Activator; DNA-binding;
KW   Trans-acting factor; DNA replication; Repressor; Nuclear protein.
SQ   SEQUENCE   368 AA;  41739 MW;  3212B423B2F629D3 CRC64;

  Query Match             49.2%;  Score 58.5;  DB 1;  Length 368;
  Best Local Similarity   61.1%;  Pred. No. 0.28;
  Matches   11;  Conservative    1;  Mismatches    3;  Indels    3;  Gaps    1;

Qy       4 NTMHYTNWTHIYI---CE 18
           ||| ||||  ||:   ||
Db     127 NTMDYTNWKEIYLLGECE 144


RESULT 9
VE2_HPV32
ID   VE2_HPV32      STANDARD;      PRT;   394 AA.
AC   P36791;
DT   01-JUN-1994 (Rel. 29, Created)
DT   01-JUN-1994 (Rel. 29, Last sequence update)
DT   15-JUL-1998 (Rel. 36, Last annotation update)
DE   Regulatory protein E2.
GN   E2.
OS   Human papillomavirus type 32.
OC   Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC   Papillomavirus.
OX   NCBI_TaxID=10612;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=94265501; PubMed=8205838;
RA   Delius H., Hofmann B.;
RT   "Primer-directed sequencing of human papillomavirus types.";
RL   Curr. Top. Microbiol. Immunol. 186:13-31(1994).
CC   -!- FUNCTION: E2 REGULATES VIRAL TRANSCRIPTION AND DNA REPLICATION.
CC       IT BINDS TO THE E2RE RESPONSE ELEMENT (5'-ACCNNNNNNGGT-3') PRESENT
CC       IN MULTIPLE COPIES IN THE REGULATORY REGION. IT CAN EITHER
CC       ACTIVATE OR REPRESS TRANSCRIPTION DEPENDING OF E2RE'S POSITION
CC       WITH REGARDS TO PROXIMAL PROMOTER ELEMENTS. REPRESSION OCCURS
CC       BY STERICALLY HINDERING THE ASSEMBLY OF THE TRANSCRIPTION
CC       INITIATION COMPLEX. THE E1-E2 COMPLEX BINDS TO THE ORIGIN OF DNA
CC       REPLICATION.
CC   -!- SUBUNIT: BINDS DNA AS A DIMER.
CC   -!- SUBCELLULAR LOCATION: Nuclear.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; X74475; CAA52552.1; -.
DR   PIR; S36512; S36512.
DR   HSSP; P17383; 1DHM.
DR   InterPro; IPR000427; E2_C.
DR   InterPro; IPR001866; E2_N.
DR   Pfam; PF00508; E2_N; 1.
DR   Pfam; PF00511; E2_C; 1.
DR   ProDom; PD000672; E2_C; 1.
DR   ProDom; PD000678; E2_N; 1.
KW   Early protein; Transcription regulation; Activator; DNA-binding;
KW   Trans-acting factor; DNA replication; Repressor; Nuclear protein.
SQ   SEQUENCE   394 AA;  45038 MW;  113C46119C2265E7 CRC64;

  Query Match             48.7%;  Score 58;  DB 1;  Length 394;
  Best Local Similarity   69.2%;  Pred. No. 0.36;
  Matches    9;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       4 NTMHYTNWTHIYI 16
           | |||| || ||:
Db     127 NAMHYTAWTFIYV 139


RESULT 10
VE2_HPV40
ID   VE2_HPV40      STANDARD;      PRT;   370 AA.
AC   P36793;
DT   01-JUN-1994 (Rel. 29, Created)
DT   01-JUN-1994 (Rel. 29, Last sequence update)
DT   15-JUL-1998 (Rel. 36, Last annotation update)
DE   Regulatory protein E2.
GN   E2.
OS   Human papillomavirus type 40.
OC   Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC   Papillomavirus.
OX   NCBI_TaxID=10615;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=94265501; PubMed=8205838;
RA   Delius H., Hofmann B.;
RT   "Primer-directed sequencing of human papillomavirus types.";
RL   Curr. Top. Microbiol. Immunol. 186:13-31(1994).
CC   -!- FUNCTION: E2 REGULATES VIRAL TRANSCRIPTION AND DNA REPLICATION.
CC       IT BINDS TO THE E2RE RESPONSE ELEMENT (5'-ACCNNNNNNGGT-3') PRESENT
CC       IN MULTIPLE COPIES IN THE REGULATORY REGION. IT CAN EITHER
CC       ACTIVATE OR REPRESS TRANSCRIPTION DEPENDING OF E2RE'S POSITION
CC       WITH REGARDS TO PROXIMAL PROMOTER ELEMENTS. REPRESSION OCCURS
CC       BY STERICALLY HINDERING THE ASSEMBLY OF THE TRANSCRIPTION
CC       INITIATION COMPLEX. THE E1-E2 COMPLEX BINDS TO THE ORIGIN OF DNA
CC       REPLICATION.
CC   -!- SUBUNIT: BINDS DNA AS A DIMER.
CC   -!- SUBCELLULAR LOCATION: Nuclear.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; X74478; CAA52570.1; -.
DR   PIR; S36558; S36558.
DR   HSSP; P17383; 1DHM.
DR   InterPro; IPR000427; E2_C.
DR   InterPro; IPR001866; E2_N.
DR   Pfam; PF00508; E2_N; 1.
DR   Pfam; PF00511; E2_C; 1.
DR   ProDom; PD000672; E2_C; 1.
DR   ProDom; PD000678; E2_N; 1.
KW   Early protein; Transcription regulation; Activator; DNA-binding;
KW   Trans-acting factor; DNA replication; Repressor; Nuclear protein.
SQ   SEQUENCE   370 AA;  41767 MW;  02C60C5137938F72 CRC64;

  Query Match             46.2%;  Score 55;  DB 1;  Length 370;
  Best Local Similarity   61.5%;  Pred. No. 0.88;
  Matches    8;  Conservative    2;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       4 NTMHYTNWTHIYI 16
           | |||| || :|:
Db     127 NAMHYTLWTTVYV 139


RESULT 11
```

```
VE2_HPV07
ID   VE2_HPV07      STANDARD;      PRT;   375 AA.
AC   P36779;
DT   01-JUN-1994 (Rel. 29, Created)
DT   01-JUN-1994 (Rel. 29, Last sequence update)
DT   15-JUL-1998 (Rel. 36, Last annotation update)
DE   Regulatory protein E2.
GN   E2.
OS   Human papillomavirus type 7.
OC   Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC   Papillomavirus.
OX   NCBI_TaxID=10620;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=94265501; PubMed=8205838;
RA   Delius H., Hofmann B.;
RT   "Primer-directed sequencing of human papillomavirus types.";
RL   Curr. Top. Microbiol. Immunol. 186:13-31(1994).
CC   -!- FUNCTION: E2 REGULATES VIRAL TRANSCRIPTION AND DNA REPLICATION.
CC       IT BINDS TO THE E2RE RESPONSE ELEMENT (5'-ACCNNNNNNGGT-3') PRESENT
CC       IN MULTIPLE COPIES IN THE REGULATORY REGION. IT CAN EITHER
CC       ACTIVATE OR REPRESS TRANSCRIPTION DEPENDING OF E2RE'S POSITION
CC       WITH REGARDS TO PROXIMAL PROMOTER ELEMENTS. REPRESSION OCCURS
CC       BY STERICALLY HINDERING THE ASSEMBLY OF THE TRANSCRIPTION
CC      .INITIATION COMPLEX. THE E1-E2 COMPLEX BINDS TO THE ORIGIN OF DNA
CC       REPLICATION.
CC   -!- SUBUNIT: BINDS DNA AS A DIMER.
CC   -!- SUBCELLULAR LOCATION: Nuclear.
CC   ---------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------------
DR   EMBL; X74463; CAA52479.1; -.
DR   PIR; S36587; S36587.
DR   HSSP; P17383; 1DHM.
DR   InterPro; IPR000427; E2_C.
DR   InterPro; IPR001866; E2_N.
DR   Pfam; PF00508; E2_N; 1.
DR   Pfam; PF00511; E2_C; 1.
DR   ProDom; PD000672; E2_C; 1.
DR   ProDom; PD000678; E2_N; 1.
KW   Early protein; Transcription regulation; Activator; DNA-binding;
KW   Trans-acting factor; DNA replication; Repressor; Nuclear protein.
SQ   SEQUENCE   375 AA;  42695 MW;  367F0583F8E984DA CRC64;

  Query Match             46.2%;  Score 55;  DB 1;  Length 375;
  Best Local Similarity   61.5%;  Pred. No. 0.9;
  Matches    8;  Conservative    2;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       4 NTMHYTNWTHIYI 16
           | |||| || :|:
Db     127 NAMHYTLWTAVYV 139


RESULT 12
VE2_HPV13
ID   VE2_HPV13      STANDARD;      PRT;   377 AA.
AC   Q02263;
DT   01-APR-1993 (Rel. 25, Created)
DT   01-APR-1993 (Rel. 25, Last sequence update)
DT   15-JUL-1998 (Rel. 36, Last annotation update)
DE   Regulatory protein E2.
GN   E2.
OS   Human papillomavirus type 13.
OC   Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC   Papillomavirus.
OX   NCBI_TaxID=10573;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=92391075; PubMed=1325697;
RA   van Ranst M., Fuse A., Fiten P., Beuken E., Pfister H., Burk R.D.,
RA   Opdenakker G.;
RT   "Human papillomavirus type 13 and pygmy chimpanzee papillomavirus
RT   type 1: comparison of the genome organizations.";
RL   Virology 190:587-596(1992).
CC   -!- FUNCTION: E2 REGULATES VIRAL TRANSCRIPTION AND DNA REPLICATION.
CC       IT BINDS TO THE E2RE RESPONSE ELEMENT (5'-ACCNNNNNNGGT-3') PRESENT
CC       IN MULTIPLE COPIES IN THE REGULATORY REGION. IT CAN EITHER
CC       ACTIVATE OR REPRESS TRANSCRIPTION DEPENDING OF E2RE'S POSITION
CC       WITH REGARDS TO PROXIMAL PROMOTER ELEMENTS. REPRESSION OCCURS
CC       BY STERICALLY HINDERING THE ASSEMBLY OF THE TRANSCRIPTION
CC       INITIATION COMPLEX. THE E1-E2 COMPLEX BINDS TO THE ORIGIN OF DNA
CC       REPLICATION.
CC   -!- SUBUNIT: BINDS DNA AS A DIMER.
CC   -!- SUBCELLULAR LOCATION: Nuclear.
CC   ---------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------------
DR   EMBL; X62843; CAA44650.1; -.
DR   PIR; D42955; W2WL13.
DR   HSSP; P17383; 1DHM.
DR   InterPro; IPR000427; E2_C.
DR   InterPro; IPR001866; E2_N.
DR   Pfam; PF00508; E2_N; 1.
DR   Pfam; PF00511; E2_C; 1.
DR   ProDom; PD000672; E2_C; 1.
DR   ProDom; PD000678; E2_N; 1.
KW   Early protein; Transcription regulation; Activator; DNA-binding;
KW   Trans-acting factor; DNA replication; Repressor; Nuclear protein.
SQ   SEQUENCE   377 AA;  43101 MW;  BE0496E486BC6303 CRC64;

  Query Match             44.1%;  Score 52.5;  DB 1;  Length 377;
  Best Local Similarity   52.9%;  Pred. No. 2;
  Matches    9;  Conservative    3;  Mismatches    2;  Indels    3;  Gaps    1;

Qy       3 CNT---MHYTNWTHIYI 16
             |||   | | :|:|::|:
Db     123 CNTDNRMDYVSWTYIYV 139


RESULT 13
VE2_HPV11
ID   VE2_HPV11      STANDARD;      PRT;   367 AA.
AC   P04015;
DT   23-OCT-1986 (Rel. 02, Created)
DT   23-OCT-1986 (Rel. 02, Last sequence update)
DT   15-JUL-1998 (Rel. 36, Last annotation update)
DE   Regulatory protein E2.
GN   E2.
OS   Human papillomavirus type 11.
OC   Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC   Papillomavirus.
OX   NCBI_TaxID=10580;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=86181601; PubMed=3008427;
RA   Dartmann K., Schwarz E., Gissmann L., Zur Hausen H.;
RT   "The nucleotide sequence and genome organization of human papilloma
RT   virus type 11.";
RL   Virology 151:124-130(1986).
CC   -!- FUNCTION: E2 REGULATES VIRAL TRANSCRIPTION AND DNA REPLICATION.
CC       IT BINDS TO THE E2RE RESPONSE ELEMENT (5'-ACCNNNNNNGGT-3') PRESENT
CC       IN MULTIPLE COPIES IN THE REGULATORY REGION. IT CAN EITHER
CC       ACTIVATE OR REPRESS TRANSCRIPTION DEPENDING OF E2RE'S POSITION
CC       WITH REGARDS TO PROXIMAL PROMOTER ELEMENTS. REPRESSION OCCURS
```

```
CC       BY STERICALLY HINDERING THE ASSEMBLY OF THE TRANSCRIPTION
CC       INITIATION COMPLEX. THE E1-E2 COMPLEX BINDS TO THE ORIGIN OF DNA
CC       REPLICATION.
CC   -!- SUBUNIT: BINDS DNA AS A DIMER.
CC   -!- SUBCELLULAR LOCATION: Nuclear.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M14119; AAA46930.1; -.
DR   PIR; A03668; W2WL11.
DR   HSSP; P17383; 1DHM.
DR   InterPro; IPR000427; E2_C.
DR   InterPro; IPR001866; E2_N.
DR   Pfam; PF00508; E2_N; 1.
DR   Pfam; PF00511; E2_C; 1.
DR   ProDom; PD000672; E2_C; 1.
DR   ProDom; PD000678; E2_N; 1.
KW   Early protein; Transcription regulation; Activator; DNA-binding;
KW   Trans-acting factor; DNA replication; Repressor; Nuclear protein.
SQ   SEQUENCE   367 AA;  41709 MW;  F7BAF74D336804F0 CRC64;

  Query Match             43.7%;  Score 52;  DB 1;  Length 367;
  Best Local Similarity   61.5%;  Pred. No. 2.3;
  Matches    8;  Conservative    1;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       4 NTMHYTNWTHIYI 16
           | | |  |||||:
Db     127 NVMEYVVWTHIYL 139


RESULT 14
VE2_RHPV1
ID   VE2_RHPV1      STANDARD;      PRT;   366 AA.
AC   P22156;
DT   01-AUG-1991 (Rel. 19, Created)
DT   01-AUG-1991 (Rel. 19, Last sequence update)
DT   15-JUL-1998 (Rel. 36, Last annotation update)
DE   Regulatory protein E2.
GN   E2.
OS   Rhesus papillomavirus type 1 (Rhpv 1).
OC   Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC   Papillomavirus.
OX   NCBI_TaxID=10570;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=91135018; PubMed=1847267;
RA   Ostrow R.S., Labresh K.V., Faras A.J.;
RT   "Characterization of the complete RhPV 1 genomic sequence and an
RT   integration locus from a metastatic tumor.";
RL   Virology 181:424-429(1991).
CC   -!- FUNCTION: E2 REGULATES VIRAL TRANSCRIPTION AND DNA REPLICATION.
CC       IT BINDS TO THE E2RE RESPONSE ELEMENT (5'-ACCNNNNNNGGT-3') PRESENT
CC       IN MULTIPLE COPIES IN THE REGULATORY REGION. IT CAN EITHER
CC       ACTIVATE OR REPRESS TRANSCRIPTION DEPENDING OF E2RE'S POSITION
CC       WITH REGARDS TO PROXIMAL PROMOTER ELEMENTS. REPRESSION OCCURS
CC       BY STERICALLY HINDERING THE ASSEMBLY OF THE TRANSCRIPTION
CC       INITIATION COMPLEX. THE E1-E2 COMPLEX BINDS TO THE ORIGIN OF DNA
CC       REPLICATION.
CC   -!- SUBUNIT: BINDS DNA AS A DIMER.
CC   -!- SUBCELLULAR LOCATION: Nuclear.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M60184; AAA79314.1; -.
DR   PIR; D38503; W2WLR1.
DR   HSSP; P17383; 1DHM.
DR   InterPro; IPR000427; E2_C.
DR   InterPro; IPR001866; E2_N.
DR   Pfam; PF00508; E2_N; 1.
DR   Pfam; PF00511; E2_C; 1.
DR   ProDom; PD000672; E2_C; 1.
DR   ProDom; PD000678; E2_N; 1.
KW   Early protein; Transcription regulation; Activator; DNA-binding;
KW   Trans-acting factor; DNA replication; Repressor; Nuclear protein.
SQ   SEQUENCE   366 AA;  41025 MW;  44A627E20DA12284 CRC64;

  Query Match             42.9%;  Score 51;  DB 1;  Length 366;
  Best Local Similarity   61.5%;  Pred. No. 3.2;
  Matches    8;  Conservative    1;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       4 NTMHYTNWTHIYI 16
           ||| |  | |||:
Db     128 NTMEYVLWGHIYV 140


RESULT 15
VE2_HPV60
ID   VE2_HPV60      STANDARD;      PRT;   404 AA.
AC   Q80944;
DT   15-JUL-1998 (Rel. 36, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   15-JUL-1998 (Rel. 36, Last annotation update)
DE   Regulatory protein E2.
GN   E2.
OS   Human papillomavirus type 60.
OC   Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC   Papillomavirus.
OX   NCBI_TaxID=40540;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Delius H.;
RL   Submitted (OCT-1995) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: E2 REGULATES VIRAL TRANSCRIPTION AND DNA REPLICATION.
CC       IT BINDS TO THE E2RE RESPONSE ELEMENT (5'-ACCNNNNNNGGT-3') PRESENT
CC       IN MULTIPLE COPIES IN THE REGULATORY REGION. IT CAN EITHER
CC       ACTIVATE OR REPRESS TRANSCRIPTION DEPENDING OF E2RE'S POSITION
CC       WITH REGARDS TO PROXIMAL PROMOTER ELEMENTS. REPRESSION OCCURS
CC       BY STERICALLY HINDERING THE ASSEMBLY OF THE TRANSCRIPTION
CC       INITIATION COMPLEX. THE E1-E2 COMPLEX BINDS TO THE ORIGIN OF DNA
CC       REPLICATION.
CC   -!- SUBUNIT: BINDS DNA AS A DIMER.
CC   -!- SUBCELLULAR LOCATION: Nuclear.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U31792; AAA79488.1; -.
DR   HSSP; P03122; 2BOP.
DR   InterPro; IPR000427; E2_C.
DR   InterPro; IPR001866; E2_N.
DR   Pfam; PF00508; E2_N; 1.
DR   Pfam; PF00511; E2_C; 1.
DR   ProDom; PD000672; E2_C; 1.
DR   ProDom; PD000678; E2_N; 1.
KW   Early protein; Transcription regulation; Activator; DNA-binding;
KW   Trans-acting factor; DNA replication; Repressor; Nuclear protein.
SQ   SEQUENCE   404 AA;  46534 MW;  73B8CAA8508720D0 CRC64;

  Query Match             42.0%;  Score 50;  DB 1;  Length 404;
```

Best Local Similarity 66.7%; Pred. No. 4.8;
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

```
Qy       4 NTMHYTNWTHIY 15
           ||  |||| :||
Db     129 NTFPYTNWEYIY 140
```

Search completed: January 14, 2003, 18:40:11
Job time : 11 secs

GenCore version 5.1.3

OM protein - protein search, using sw model

Run on: January 14, 2003, 18:37:13 ; Search time 36 Seconds
(without alignments)
70.327 Million cell updates/sec

Title: US-09-828-645-1
Perfect score: 119
Sequence: 1 DICNTMHYTNWTHIYICEE 19

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_101002:*
1: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1982.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1983.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1984.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1985.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1986.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1987.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1988.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1989.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1990.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1991.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1992.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1993.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1994.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1995.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1996.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1997.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1998.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1999.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA2000.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA2001.DAT:*
23: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	119	100.0	19	23	AAU10801	Human papillomavir
2	119	100.0	20	12	AAR15550	Immunopeptide #1 d
3	119	100.0	147	23	AAE15402	HPV 16 E2 protein.
4	119	100.0	201	22	AAB74333	Human papillomavir
5	119	100.0	365	18	AAW12095	Human papillomavir
6	119	100.0	365	19	AAW50005	Human papillomavir
7	119	100.0	365	22	AAB98418	Human papillomavir
8	119	100.0	365	22	AAB35191	Human papillomavir
9	119	100.0	365	22	AAB35192	Human papillomavir
10	119	100.0	365	23	AAU10809	Human papillomavir
11	109	91.6	20	12	AAR15553	Immunopeptide #4 d
12	97	81.5	19	23	AAU10806	Human papillomavir
13	79	66.4	20	12	AAR15591	Immunopeptide #1 d
14	79	66.4	372	22	AAB98432	Human papillomavir
15	66	55.5	353	22	AAB98445	Human papillomavir
16	57	47.9	20	12	AAR15596	Immunopeptide deri
17	52	43.7	20	12	AAR15535	Immunopeptide deri
18	52	43.7	148	23	AAE15404	HPV 11 E2 protein.
19	52	43.7	201	22	AAB74335	Human papillomavir
20	52	43.7	315	23	AAE18268	Human G-protein co
21	52	43.7	315	23	AAE18272	Human G-protein co
22	52	43.7	315	23	AAE18273	Human G-protein co
23	52	43.7	315	23	AAE18274	Human G-protein co
24	52	43.7	315	23	AAE18275	Human G-protein co
25	52	43.7	315	23	AAE18276	Human G-protein co
26	52	43.7	315	23	AAE18277	Human G-protein co
27	52	43.7	315	23	AAE18278	Human G-protein co
28	52	43.7	315	23	AAE18279	Human G-protein co
29	52	43.7	315	23	AAE18280	Human G-protein co
30	52	43.7	317	23	AAU95677	Human olfactory an
31	52	43.7	334	22	AAU24698	Human olfactory re
32	52	43.7	334	22	AAG71887	Human olfactory re
33	52	43.7	334	22	AAG67468	Breast amplified G
34	52	43.7	334	23	ABB81441	Human GPCR2a prote
35	52	43.7	334	23	ABB81442	Human GPCR2b prote
36	52	43.7	334	23	ABB81443	Human GPCR2c prote
37	52	43.7	334	23	ABB81444	Human GPCR2d prote
38	52	43.7	334	23	AAU85318	G-coupled olfactor
39	52	43.7	367	22	AAB98409	Human papillomavir
40	52	43.7	367	23	ABB07666	Mutated HPV11 E2 p
41	52	43.7	367	23	ABB07667	Mutated HPV11 E2 p
42	52	43.7	367	23	ABB07668	Wild-type HPV11 E2
43	50	42.0	363	22	AAM78368	Human protein SEQ
44	50	42.0	436	21	AAB57062	Human prostate can
45	50	42.0	455	20	AAW78245	Fragment of human

ALIGNMENTS

RESULT 1
AAU10801
ID AAU10801 standard; peptide; 19 AA.
XX
AC AAU10801;
XX
DT 14-FEB-2002 (first entry)
XX
DE Human papillomavirus (HPV) 16 E2 coding region derived peptide #1.
XX
KW Human papillomavirus 16; HPV 16; cancer; squamous cell carcinoma;
KW adenocarcinoma; koilocytosis; hyperkerotosis; intraepithelial neoplasia;
KW intraepithelial lesion; dysplasia; head cancer; neck cancer;
KW small cell lung cancer; melanoma; oncogene.
XX
OS Human papillomavirus 16.
XX
PN WO200177142-A1.
XX
PD 18-OCT-2001.
XX
PF 05-APR-2001; 2001WO-US11233.
XX
PR 05-APR-2000; 2000US-194796P.
XX
PA (IMPA-) IMPACT DIAGNOSTICS INC.
XX
PI Hu YX;
XX
DR WPI; 2002-010888/01.
XX
PT New peptides derived from E2, E6 or E7 early coding regions of human
PT papillomavirus 16 and 18, useful in diagnosis of human papillomavirus

PT infection and associated malignancy e.g. cervical carcinoma -
XX
PS Claim 2; Fig 2; 28pp; English.
XX
CC The invention describes a novel peptide derived from the E2, E6 or E7
CC early coding region of human papillomavirus (HPV) 16 and 18, which is
CC soluble in aqueous solution and has a lysine or cysteine residue near the
CC amino terminus, very few tryptophan, methionine and cysteine residues,
CC and/or many glycine and asparagine residues. The peptides and diagnostic
CC method are used to diagnose HPV infection, especially infection with
CC oncogenic HPV by using peptides derived from the E2 region, since HPV 16
CC and 18 are the main HPV genetic types associated with cancers, and
CC presence of antibodies to E2 protein is known to provide evidence of HPV
CC infection. They are also useful to diagnose HPV associated malignancy or
CC premalignancy, especially carcinoma by using peptides derived from the E6
CC or E7 regions, since E6 and E7 are thought to be tumour-specific
CC antigens. The peptides and diagnostic method are especially useful to
CC diagnose cervical carcinoma (e.g. adenocarcinoma of the uterine cervix)
CC and any HPV associated epithelial cell abnormality including high grade
CC dysplasias, koilocytosis, hyperkerotosis, precancerous conditions
CC encompassing intraepithelial neoplasias or intraepithelial lesions, and
CC invasive or malignant cancers. They are also used to detect head and neck
CC cancers, small cell lung cancers, penal and anal squamous cell carcinomas
CC and melanoma. This is the amino acid sequence of peptide epitope #1,
CC derived from the E2 early coding region of HPV 16, an oncoprotein
CC involved in the integration of the viral genome into the host cell genome
CC by the HPV types associated with malignant cancers, described in the
CC method of the invention.
XX
SQ Sequence 19 AA;

Query Match 100.0%; Score 119; DB 23; Length 19;
Best Local Similarity 100.0%; Pred. No. 1.1e-10;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DICNTMHYTNWTHIYICEE 19
||||||||||||||||||
Db 1 DICNTMHYTNWTHIYICEE 19

RESULT 2
AAR15550
ID AAR15550 standard; Protein; 20 AA.
XX
AC AAR15550;
XX
DT 02-MAR-1992 (first entry)
XX
DE Immunopeptide #1 derived from HPV16 E2 peptide.
XX
KW cervical cancer; cervical intraepithelial neoplasia; CIN; wart;
KW squamous cell carcinoma; ELISA; HPV 16.
XX
OS Synthetic.
XX
PN WO9118294-A.
XX
PD 28-NOV-1991.
XX
PF 13-MAY-1991; 91WO-SE00335.
XX
PR 11-MAY-1990; 90SE-0001705.
XX
PA (MEDS-) MEDSCAND AB.
XX
PI Dillner J, Dillner L, Cheng HM;
XX
DR WPI; 1991-369390/50.
XX
PT Diagnosis of human papilloma virus infection and PV-carrying
PT tumours - using synthetic peptide(s) to detect virus specific
PT antigen-antibody complexes by immunoassay
XX
PS Disclosure; Page 38; 72pp; English.
XX
CC This is one of a large number of peptides which have been
CC synthesised on the basis of the amino acid sequences for the E2, E4,
CC E7, L1 or L2 proteins of HPV 1, 5, 6, 8, 11, 16, 18, 31 and 33. The
CC selection of peptide sequences was based on the assumption that an
CC immunoreactive region might be situated in the same relative region
CC of a protein from different HPV types. The peptides were used in
CC diagnostic immunoassays to detect HPV-infection.
CC See AAR15523-R15601.
XX
SQ Sequence 20 AA;

Query Match 100.0%; Score 119; DB 12; Length 20;
Best Local Similarity 100.0%; Pred. No. 1.2e-10;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DICNTMHYTNWTHIYICEE 19
||||||||||||||||||
Db 2 DICNTMHYTNWTHIYICEE 20

RESULT 3
AAE15402
ID AAE15402 standard; Protein; 147 AA.
XX
AC AAE15402;
XX
DT 26-MAR-2002 (first entry)
XX
DE HPV 16 E2 protein.
XX
KW Human papillomavirus; HPV 16 E2 protein; cytotoxic gene; anogenital SCC;
KW cancer; intraepithelial neoplasia; IN; squamous cell carcinoma;
KW tumour; gene therapy; keratinocyte.
XX
OS Human papillomavirus.
XX
PN WO200187350-A2.
XX
PD 22-NOV-2001.
XX
PF 11-MAY-2001; 2001WO-US15407.
XX
PR 12-MAY-2000; 2000US-203709P.
XX
PA (REGC) UNIV CALIFORNIA.
XX
PI Sethi N, Palefsky J;
XX
DR WPI; 2002-082947/11.
XX
PT Novel gene therapy approach to specifically eliminate keratinocytes or
PT other cells expressing human papilloma virus, involves transfecting
PT cell with a construct encoding HPV specific promoter induced by a HPV
PT protein -
XX
PS Example 1; Fig 13; 72pp; English.
XX
CC The invention provides a novel gene therapy approach to specifically
CC eliminate keratinocytes or other cells expressing early
CC human papillomavirus (HPV) and which is minimally toxic to HPV-negative
CC cells. The method involves transfecting a mammalian cell with a
CC nucleic acid construct encoding a HPV specific promoter that is induced
CC by a HPV protein where the promoter is operably linked to a nucleic
CC acid comprising a cytotoxic gene such that the cell, when infected with
CC a HPV, induces expression of the cytotoxic gene resulting in death of
CC the mammalian cell. The method is useful for selectively killing a
CC cancer cell comprising intraepithelial neoplasia (IN), anogenital
CC cancer or a metastatic cell or solid tumour bearing a HPV. The nucleic
CC acid construct comprising HPV promoter operably linked to a reporter
CC gene is useful for selectively labelling a cell bearing HPV. The method
CC is useful for specifically eliminating keratinocytes, particularly cells

```
CC   of anogenital squamous cell carcinoma (SCC), or other cells expressing
CC   early HPV e.g. HPV 16 genes and a cell comprising a wart. The present
CC   sequence is HPV 16 E2 protein.
XX
SQ   Sequence   147 AA;

  Query Match            100.0%;  Score 119;  DB 23;  Length 147;
  Best Local Similarity  100.0%;  Pred. No. 9.9e-10;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 DICNTMHYTNWTHIYICEE 19
          ||||||||||||||||||||
Db     60 DICNTMHYTNWTHIYICEE 78


RESULT 4
AAB74333
ID   AAB74333 standard; Protein; 201 AA.
XX
AC   AAB74333;
XX
DT   29-JUN-2001  (first entry)
XX
DE   Human papillomavirus 16 E2NT module.
XX
KW   E2NT; antiviral; warts; proliferative skin lesion; cervical cancer;
KW   HPV.
XX
OS   Human papillomavirus.
XX
PN   WO200121645-A2.
XX
PD   29-MAR-2001.
XX
PF   18-SEP-2000; 2000WO-GB03568.
XX
PR   17-SEP-1999;   99GB-0021938.
XX
PA   (UYYO-) UNIV YORK.
XX
PI   Antson A,  Maitland N;
XX
DR   WPI; 2001-328091/34.
XX
PT   Crystallized molecular complex of the nuclear phosphoprotein E2
PT   N-terminal module useful in drug design for the treatment of conditions
PT   associated with human papilloma virus infection  -
XX
PS   Disclosure; Fig 1; 87pp; English.
XX
CC   The present invention relates to a crystallized molecular complex
CC   of an E2 N-terminal module (E2NT) dimer protein or its homologue having
CC   residues vital for transcription and replication activities of the
CC   protein. The invention is used to identify antiviral agents, and to
CC   prepare treatments for conditions that arise from herpes papillomavirus
CC   infection, particularly warts, proliferative skin lesions and/or
CC   cervical cancer. The present sequence is the E2NT module of HPV16.
XX
SQ   Sequence   201 AA;

  Query Match            100.0%;  Score 119;  DB 22;  Length 201;
  Best Local Similarity  100.0%;  Pred. No. 1.4e-09;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 DICNTMHYTNWTHIYICEE 19
          ||||||||||||||||||||
Db    124 DICNTMHYTNWTHIYICEE 142


RESULT 5
AAW12095
ID   AAW12095 standard; Protein; 365 AA.
XX
AC   AAW12095;
XX
DT   28-APR-1997  (first entry)
XX
DE   Human papillomavirus-16 E2 protein.
XX
KW   Papillomavirus E1/E2 interaction domain; cancer; diagnosis; BPV-1;
KW   HPV-16; E2 protein.
XX
OS   Human papillomavirus type 16.
XX
FH   Key             Location/Qualifiers
FT   Domain          1..190
FT                   /label= E1/E2_interaction domain
FT                   /note= "E1/E2 interaction domain preferred for use
FT                          in methods for the invention"
XX
PN   WO9641018-A1.
XX
PD   19-DEC-1996.
XX
PF   24-MAY-1996;   96WO-US07615.
XX
PR   07-JUN-1995;   95US-0472666.
XX
PA   (HARD ) HARVARD COLLEGE.
XX
PI   Benson JD,  Howley PM,  Sakai H,  Yasugi T;
XX
DR   WPI; 1997-077280/07.
XX
PT   Determining pathogenic state of papilloma virus (PV) infected cell -
PT   by testing ability to bind PV E1/E2 interaction domain, failure to
PT   bind indicating a (pre)cancerous condition
XX
PS   Claim 2; Page 54-55; 72pp; English.
XX
CC   The human papillomavirus (HPV) E2 protein (AAW12095) contains an
CC   E1/E2 interaction domain that specifically reacts with the HPV E1
CC   protein (AAW12094).  The interaction domain was identified using a
CC   yeast interaction trap system with E1 fragments as bait and E2
CC   fragments as prey.  Bovine papillomavirus (HPV) E1/E2 interaction
CC   domains have also been identified (see AAW12092-93).  HPV and BPV
CC   E1/E2 interaction domains can be used to determine the pathogenic
CC   status of a cell; non-cancerous PV-infected cells will bind the
CC   interaction domain, but (pre)cancerous cells will not.  The
CC   interaction domains can also be used in methods for typing HPV
CC   infections, inhibiting PV replication, and identifying inhibitors
CC   of PV replication.
XX
SQ   Sequence   365 AA;

  Query Match            100.0%;  Score 119;  DB 18;  Length 365;
  Best Local Similarity  100.0%;  Pred. No. 2.6e-09;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 DICNTMHYTNWTHIYICEE 19
          ||||||||||||||||||||
Db    124 DICNTMHYTNWTHIYICEE 142


RESULT 6
AAW50005
ID   AAW50005 standard; Protein; 365 AA.
XX
AC   AAW50005;
XX
DT   11-JUN-1998  (first entry)
XX
DE   Human papillomavirus-16 E2 protein.
XX
KW   Human papillomavirus-16; HPV-16; E2 protein; infection; treatment;
KW   wart; epidermodysplasia verruciformis; laryngeal papilloma;
```

```
KW   cervical carcinoma.
XX
OS   Human papillomavirus.
XX
FH   Key             Location/Qualifiers
FT   Domain          1..198
FT                   /note= "transcription activating domain"
FT   Domain          281..365
FT                   /note= "DNA-binding domain"
XX
PN   WO9801148-A1.
XX
PD   15-JAN-1998.
XX
PF   08-JUL-1997;   97WO-US11815.
XX
PR   09-JUL-1996;   96US-0677206.
XX
PA   (HARD ) HARVARD COLLEGE.
XX
PI   Dowhanick JJ,  Howley PM;
XX
DR   WPI; 1998-100815/09.
DR   N-PSDB; AAV18817.
XX
PT   Treating papilloma virus infections, by repressing E6 and E7
PT   oncogenic proteins - using E2 polypeptide or nucleic acid encoding
PT   it, useful for treating warts and cervical cancer
XX
PS   Claim 22; Pages 48-49; 71pp; English.
XX
CC   The present sequence is the human papillomavirus-16 (HPV-16)
CC   E2 protein.
CC   Papillomavirus infections can be treated by administering an E2
CC   protein, comprising a DNA-binding and transcription activating
CC   domain, or the gene encoding the E2 protein to infected cells. The
CC   method is applicable to livestock, zoo animals, pets and humans,
CC   specifically for the treatment, prevention or reversal of warts,
CC   e.g. plantar, common, Butcher's common, flat or genital warts, or
CC   epidermodysplasia verruciformis, or growth of papillomavirus
CC   transformed cells, especially laryngeal papilloma or cervical
CC   carcinoma.
XX
SQ   Sequence   365 AA;

  Query Match             100.0%;  Score 119;  DB 19;  Length 365;
  Best Local Similarity   100.0%;  Pred. No. 2.6e-09;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 DICNTMHYTNWTHIYICEE 19
          ||||||||||||||||||||
Db    124 DICNTMHYTNWTHIYICEE 142


RESULT 7
AAB98418
ID   AAB98418 standard; Protein; 365 AA.
XX
AC   AAB98418;
XX
DT   22-AUG-2001  (first entry)
XX
DE   Human papillomavirus protein HPV16 E2.
XX
KW   Human papillomavirus; human leukocyte antigen; HLA; immune response;
KW   HPV; epitope; T cell; identification; vaccine; infection; genital wart;
KW   neoplastic growth; antiviral.
XX
OS   Human papillomavirus.
XX
PN   WO200141799-A1.
XX
PD   14-JUN-2001.
XX
PF   11-DEC-2000; 2000WO-US33549.
XX
PR   10-DEC-1999;   99US-0172705.
PR   15-AUG-2000; 2000US-0641528.
XX
PA   (EPIM-) EPIMMUNE INC.
XX
PI   Sette A,  Sidney J,  Southwood S,  Chesnut R,  Celis E,  Grey HM;
XX
DR   WPI; 2001-381497/40.
XX
PT   An isolated human papilloma virus (HPV) epitope, useful in vaccines for
PT   treating HPV infections -
XX
PS   Disclosure; Page 20; 756pp; English.
XX
CC   The present invention describes an isolated prepared human papillomavirus
CC   (HPV) epitope (I). (I) has antiviral activity, and can be used in
CC   vaccine production. Peptides and corresponding nucleic acid compositions
CC   from the present invention are useful for stimulating an immune response
CC   to HPV by stimulating the production of CTL or HTL responses,
CC   specifically in the treatment or prophylaxis of HPV infection, in persons
CC   who have not manifested symptoms e.g. genital warts or neoplastic growth.
CC   The peptides can also be used in a tetramer staining assay to assess
CC   peripheral blood mononuclear cells for the presence of antigen-specific
CC   CTLs following exposure to a pathogen or immunogen, and as reagents to
CC   evaluate immune recall responses or evaluate the efficacy of a vaccine.
CC   The vaccine compositions are useful for removing warts or treating HPV
CC   infections. The epitopes for inclusion in an epitope-base vaccine may
CC   be selected from conserved regions of viral or tumour-associated
CC   antigens, which reduces the likelihood of escape mutants, also
CC   immunosuppressive epitopes that may be present in whole antigens can be
CC   avoided with the use of epitope-base vaccines. An additional advantage
CC   is the ability to combine selected epitopes (CTL and HTL) and to modify
CC   the composition of the epitopes achieving enhanced immunogenicity, the
CC   major benefit of the vaccine is that is safe and efficacious. AAB98391
CC   to AAB98477 represent polypeptide sequences used in the exemplification
CC   of the present invention.
XX
SQ   Sequence   365 AA;

  Query Match             100.0%;  Score 119;  DB 22;  Length 365;
  Best Local Similarity   100.0%;  Pred. No. 2.6e-09;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 DICNTMHYTNWTHIYICEE 19
          ||||||||||||||||||||
Db    124 DICNTMHYTNWTHIYICEE 142


RESULT 8
AAB35191
ID   AAB35191 standard; Protein; 365 AA.
XX
AC   AAB35191;
XX
DT   24-APR-2001  (first entry)
XX
DE   Human papillomavirus HPV16 E2 protein.
XX
KW   Papillomavirus; apoptosis; cell death induction; E2 protein; E7 protein;
KW   p53; cancer.
XX
OS   Human papillomavirus.
XX
PN   WO200103717-A2.
XX
PD   18-JAN-2001.
XX
PF   13-JUL-2000; 2000WO-GB02693.
XX
PR   13-JUL-1999;   99GB-0016363.
```

```
XX
PA   (UYBR-) UNIV BRISTOL.
XX
PI   Gaston KL,  Stern PL,  Clarke AR;
XX
DR   WPI; 2001-123139/13.
DR   N-PSDB; AAF24352.
XX
PT   Inducing cell death in papillomavirus (PV) positive or negative
PT   oncogenic cells, wild-type, p53 mutant- or p53 related gene positive
PT   cells using PV E2 and/or E7 protein or its functional portion or
PT   derivative -
XX
PS   Disclosure; Fig 9; 55pp; English.
XX
CC   The present invention describes a method of inducing apoptotic cell death
CC   in papillomavirus negative and positive cells using the papillomavirus
CC   E2 and E7 proteins. This can be used in the treatment of epithelial cell
CC   cancers, including cervical cancer. The present sequence is the HPV16
CC   E2 protein.
XX
SQ   Sequence   365 AA;

  Query Match             100.0%;  Score 119;  DB 22;  Length 365;
  Best Local Similarity   100.0%;  Pred. No. 2.6e-09;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 DICNTMHYTNWTHIYICEE 19
           ||||||||||||||||||||
Db     124 DICNTMHYTNWTHIYICEE 142


RESULT 9
AAB35192
ID   AAB35192 standard; Protein; 365 AA.
XX
AC   AAB35192;
XX
DT   24-APR-2001  (first entry)
XX
DE   Human papillomavirus HPV16 E2DBM protein.
XX
KW   Papillomavirus; apoptosis; cell death induction; E2 protein; E7 protein;
KW   p53; cancer.
XX
OS   Human papillomavirus.
XX
PN   WO200103717-A2.
XX
PD   18-JAN-2001.
XX
PF   13-JUL-2000; 2000WO-GB02693.
XX
PR   13-JUL-1999;   99GB-0016363.
XX
PA   (UYBR-) UNIV BRISTOL.
XX
PI   Gaston KL,  Stern PL,  Clarke AR;
XX
DR   WPI; 2001-123139/13.
DR   N-PSDB; AAF24353.
XX
PT   Inducing cell death in papillomavirus (PV) positive or negative
PT   oncogenic cells, wild-type, p53 mutant- or p53 related gene positive
PT   cells using PV E2 and/or E7 protein or its functional portion or
PT   derivative -
XX
PS   Disclosure; Fig 10; 55pp; English.
XX
CC   The present invention describes a method of inducing apoptotic cell death
CC   in papillomavirus negative and positive cells using the papillomavirus
CC   E2 and E7 proteins. This can be used in the treatment of epithelial cell
CC   cancers, including cervical cancer. The present sequence is the HPV16
CC   E2DBM protein.
XX
SQ   Sequence   365 AA;

  Query Match             100.0%;  Score 119;  DB 22;  Length 365;
  Best Local Similarity   100.0%;  Pred. No. 2.6e-09;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 DICNTMHYTNWTHIYICEE 19
           ||||||||||||||||||||
Db     124 DICNTMHYTNWTHIYICEE 142


RESULT 10
AAU10809
ID   AAU10809 standard; Protein; 365 AA.
XX
AC   AAU10809;
XX
DT   14-FEB-2002  (first entry)
XX
DE   Human papillomavirus (HPV) 16, E2 coding region.
XX
KW   Human papillomavirus 16; HPV 16; cancer; squamous cell carcinoma;
KW   adenocarcinoma; koilocytosis; hyperkerotosis; intraepithelial neoplasia;
KW   intraepithelial lesion; dysplasia; head cancer; neck cancer;
KW   small cell lung cancer; melanoma; oncogene.
XX
OS   Human papillomavirus 16.
XX
FH   Key             Location/Qualifiers
FT   Peptide         124..142
FT                   /label= E2_peptide_epitope_1
FT                   /note= "This region is specifically referred to
FT                         in claim 2"
FT   Peptide         328..343
FT                   /label= E2_peptide_epitope_2
FT                   /note= "This region is specifically referred to
FT                         in claim 2"
XX
PN   WO200177142-A1.
XX
PD   18-OCT-2001.
XX
PF   05-APR-2001; 2001WO-US11233.
XX
PR   05-APR-2000; 2000US-194796P.
XX
PA   (IMPA-) IMPACT DIAGNOSTICS INC.
XX
PI   Hu YX;
XX
DR   WPI; 2002-010888/01.
XX
PT   New peptides derived from E2, E6 or E7 early coding regions of human
PT   papillomavirus 16 and 18, useful in diagnosis of human papillomavirus
PT   infection and associated malignancy e.g. cervical carcinoma -
XX
PS   Disclosure; Fig 2; 28pp; English.
XX
CC   The invention describes a novel peptide derived from the E2, E6 or E7
CC   early coding region of human papillomavirus (HPV) 16 and 18, which is
CC   soluble in aqueous solution and has a lysine or cysteine residue near the
CC   amino terminus, very few tryptophan, methionine and cysteine residues,
CC   and/or many glycine and asparagine residues. The peptides and diagnostic
CC   method are used to diagnose HPV infection, especially infection with
CC   oncogenic HPV by using peptides derived from the E2 region, since HPV 16
CC   and 18 are the main HPV genetic types associated with cancers, and
CC   presence of antibodies to E2 protein is known to provide evidence of HPV
CC   infection. They are also useful to diagnose HPV associated malignancy or
CC   premalignancy, especially carcinoma by using peptides derived from the E6
CC   or E7 regions, since E6 and E7 are thought to be tumour-specific
CC   antigens. The peptides and diagnostic method are especially useful to
```

```
CC   diagnose cervical carcinoma (e.g. adenocarcinoma of the uterine cervix)
CC   and any HPV associated epithelial cell abnormality including high grade
CC   dysplasias, koilocytosis, hyperkerotosis, precancerous conditions
CC   encompassing intraepithelial neoplasias or intraepithelial lesions, and
CC   invasive or malignant cancers. They are also used to detect head and neck
CC   cancers, small cell lung cancers, penal and anal squamous cell carcinomas
CC   and melanoma. This is the amino acid sequence of the E2 early coding
CC   region of HPV 16, an oncoprotein involved in the integration of the viral
CC   genome into the host cell genome by the HPV types associated with
CC   malignant cancers, described in the method of the invention.
XX
SQ   Sequence   365 AA;

  Query Match             100.0%;  Score 119;  DB 23;  Length 365;
  Best Local Similarity   100.0%;  Pred. No. 2.6e-09;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 DICNTMHYTNWTHIYICEE 19
          |||||||||||||||||||
Db    124 DICNTMHYTNWTHIYICEE 142


RESULT 11
AAR15553
ID   AAR15553 standard; Protein; 20 AA.
XX
AC   AAR15553;
XX
DT   02-MAR-1992  (first entry)
XX
DE   Immunopeptide #4 derived from HPV16 E2 peptide.
XX
KW   cervical cancer; cervical intraepithelial neoplasia; CIN; wart;
KW   squamous cell carcinoma; ELISA; HPV 16.
XX
OS   Synthetic.
XX
PN   WO9118294-A.
XX
PD   28-NOV-1991.
XX
PF   13-MAY-1991;   91WO-SE00335.
XX
PR   11-MAY-1990;   90SE-0001705.
XX
PA   (MEDS-) MEDSCAND AB.
XX
PI   Dillner J,  Dillner L,  Cheng HM;
XX
DR   WPI; 1991-369390/50.
XX
PT   Diagnosis of human papilloma virus infection and PV-carrying
PT   tumours - using synthetic peptide(s) to detect virus specific
PT   antigen-antibody complexes by immunoassay
XX
PS   Disclosure; Page 38; 72pp; English.
XX
CC   This is one of a large number of peptides which have been
CC   synthesised on the basis of the amino acid sequences for the E2, E4,
CC   E7, L1 or L2 proteins of HPV 1, 5, 6, 8, 11, 16, 18, 31 and 33. The
CC   selection of peptide sequences was based on the assumption that an
CC   immunoreactive region might be situated in the same relative region
CC   of a protein from different HPV types. The peptides were used in
CC   diagnostic immunoassays to detect HPV-infection.
CC   See AAR15523-R15601.
XX
SQ   Sequence   20 AA;

  Query Match             91.6%;  Score 109;  DB 12;  Length 20;
  Best Local Similarity   100.0%;  Pred. No. 3.4e-09;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 DICNTMHYTNWTHIYIC 17
          |||||||||||||||||
Db      4 DICNTMHYTNWTHIYIC 20


RESULT 12
AAU10806
ID   AAU10806 standard; peptide; 19 AA.
XX
AC   AAU10806;
XX
DT   14-FEB-2002  (first entry)
XX
DE   Human papillomavirus (HPV) 16 E2 coding region derived peptide #3.
XX
KW   Human papillomavirus 16; HPV 16; cancer; squamous cell carcinoma;
KW   adenocarcinoma; koilocytosis; hyperkerotosis; intraepithelial neoplasia;
KW   intraepithelial lesion; dysplasia; head cancer; neck cancer;
KW   small cell lung cancer; melanoma; oncogene.
XX
OS   Human papillomavirus 16.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   3
FT                   /label= OTHER
FT                   /note= "OTHER= L-carboxymethylcysteine"
FT   Modified-site   17
FT                   /label= OTHER
FT                   /note= "OTHER= L-carboxymethylcysteine"
XX
PN   WO200177142-A1.
XX
PD   18-OCT-2001.
XX
PF   05-APR-2001; 2001WO-US11233.
XX
PR   05-APR-2000; 2000US-194796P.
XX
PA   (IMPA-) IMPACT DIAGNOSTICS INC.
XX
PI   Hu YX;
XX
DR   WPI; 2002-010888/01.
XX
PT   New peptides derived from E2, E6 or E7 early coding regions of human
PT   papillomavirus 16 and 18, useful in diagnosis of human papillomavirus
PT   infection and associated malignancy e.g. cervical carcinoma  -
XX
PS   Disclosure; Page 25; 28pp; English.
XX
CC   The invention describes a novel peptide derived from the E2, E6 or E7
CC   early coding region of human papillomavirus (HPV) 16 and 18, which is
CC   soluble in aqueous solution and has a lysine or cysteine residue near the
CC   amino terminus, very few tryptophan, methionine and cysteine residues,
CC   and/or many glycine and asparagine residues. The peptides and diagnostic
CC   method are used to diagnose HPV infection, especially infection with
CC   oncogenic HPV by using peptides derived from the E2 region, since HPV 16
CC   and 18 are the main HPV genetic types associated with cancers, and
CC   presence of antibodies to E2 protein is known to provide evidence of HPV
CC   infection. They are also useful to diagnose HPV associated malignancy or
CC   premalignancy, especially carcinoma by using peptides derived from the E6
CC   or E7 regions, since E6 and E7 are thought to be tumour-specific
CC   antigens. The peptides and diagnostic method are especially useful to
CC   diagnose cervical carcinoma (e.g. adenocarcinoma of the uterine cervix)
CC   and any HPV associated epithelial cell abnormality including high grade
CC   dysplasias, koilocytosis, hyperkerotosis, precancerous conditions
CC   encompassing intraepithelial neoplasias or intraepithelial lesions, and
CC   invasive or malignant cancers. They are also used to detect head and neck
CC   cancers, small cell lung cancers, penal and anal squamous cell carcinomas
CC   and melanoma. This is the amino acid sequence of peptide epitope #3, a
CC   synthetic peptide derived from the E2 early coding region of HPV 16, an
CC   oncoprotein involved in the integration of the viral genome into the host
CC   cell genome by the HPV types associated with malignant cancers, described
```

```
CC   in the method of the invention.
XX
SQ   Sequence   19 AA;

  Query Match            81.5%;  Score 97;  DB 23;  Length 19;
  Best Local Similarity  89.5%;  Pred. No. 1.8e-07;
  Matches   17;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 DICNTMHYTNWTHIYICEE 19
          || |||||||||||| ||
Db      1 DIXNTMHYTNWTHIYIXEE 19


RESULT 13
AAR15591
ID   AAR15591 standard; Protein; 20 AA.
XX
AC   AAR15591;
XX
DT   02-MAR-1992  (first entry)
XX
DE   Immunopeptide #1 derived from HPV31 E2 peptide.
XX
KW   cervical cancer; cervical intraepithelial neoplasia; CIN; wart;
KW   squamous cell carcinoma; ELISA; HPV 31.
XX
OS   Synthetic.
XX
PN   WO9118294-A.
XX
PD   28-NOV-1991.
XX
PF   13-MAY-1991;   91WO-SE00335.
XX
PR   11-MAY-1990;   90SE-0001705.
XX
PA   (MEDS-) MEDSCAND AB.
XX
PI   Dillner J,  Dillner L,  Cheng HM;
XX
DR   WPI; 1991-369390/50.
XX
PT   Diagnosis of human papilloma virus infection and PV-carrying
PT   tumours - using synthetic peptide(s) to detect virus specific
PT   antigen-antibody complexes by immunoassay
XX
PS   Disclosure; Page 39; 72pp; English.
XX
CC   This is one of a large number of peptides which have been
CC   synthesised on the basis of the amino acid sequences for the E2, E4,
CC   E7, L1 or L2 proteins of HPV 1, 5, 6, 8, 11, 16, 18, 31 and 33. The
CC   selection of peptide sequences was based on the assumption that an
CC   immunoreactive region might be situated in the same relative region
CC   of a protein from different HPV types. The peptides were used in
CC   diagnostic immunoassays to detect HPV-infection.
CC   See AAR15523-R15601.
XX
SQ   Sequence   20 AA;

  Query Match            66.4%;  Score 79;  DB 12;  Length 20;
  Best Local Similarity  70.6%;  Pred. No. 7.6e-05;
  Matches   12;  Conservative    2;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      1 DICNTMHYTNWTHIYIC 17
          |: |||||||  ||:|
Db      4 DVHNTMHYTNWKFIYLC 20


RESULT 14
AAB98432
ID   AAB98432 standard; Protein; 372 AA.
XX
AC   AAB98432;
XX
DT   22-AUG-2001  (first entry)
XX
DE   Human papillomavirus protein HPV31 E2.
XX
KW   Human papillomavirus; human leukocyte antigen; HLA; immune response;
KW   HPV; epitope; T cell; identification; vaccine; infection; genital wart;
KW   neoplastic growth; antiviral.
XX
OS   Human papillomavirus.
XX
PN   WO200141799-A1.
XX
PD   14-JUN-2001.
XX
PF   11-DEC-2000; 2000WO-US33549.
XX
PR   10-DEC-1999;   99US-0172705.
PR   15-AUG-2000; 2000US-0641528.
XX
PA   (EPIM-) EPIMMUNE INC.
XX
PI   Sette A,  Sidney J,  Southwood S,  Chesnut R,  Celis E,  Grey HM;
XX
DR   WPI; 2001-381497/40.
XX
PT   An isolated human papilloma virus (HPV) epitope, useful in vaccines for
PT   treating HPV infections -
XX
PS   Disclosure; Page 23; 756pp; English.
XX
CC   The present invention describes an isolated prepared human papillomavirus
CC   (HPV) epitope (I). (I) has antiviral activity, and can be used in
CC   vaccine production. Peptides and corresponding nucleic acid compositions
CC   from the present invention are useful for stimulating an immune response
CC   to HPV by stimulating the production of CTL or HTL responses,
CC   specifically in the treatment or prophylaxis of HPV infection, in persons
CC   who have not manifested symptoms e.g. genital warts or neoplastic growth.
CC   The peptides can also be used in a tetramer staining assay to assess
CC   peripheral blood mononuclear cells for the presence of antigen-specific
CC   CTLs following exposure to a pathogen or immunogen, and as reagents to
CC   evaluate immune recall responses or evaluate the efficacy of a vaccine.
CC   The vaccine compositions are useful for removing warts or treating HPV
CC   infections. The epitopes for inclusion in an epitope-base vaccine may
CC   be selected from conserved regions of viral or tumour-associated
CC   antigens, which reduces the likelihood of escape mutants, also
CC   immunosuppressive epitopes that may be present in whole antigens can be
CC   avoided with the use of epitope-base vaccines. An additional advantage
CC   is the ability to combine selected epitopes (CTL and HTL) and to modify
CC   the composition of the epitopes achieving enhanced immunogenicity, the
CC   major benefit of the vaccine is that is safe and efficacious. AAB98391
CC   to AAB98477 represent polypeptide sequences used in the exemplification
CC   of the present invention.
XX
SQ   Sequence   372 AA;

  Query Match            66.4%;  Score 79;  DB 22;  Length 372;
  Best Local Similarity  70.6%;  Pred. No. 0.0017;
  Matches   12;  Conservative    2;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      1 DICNTMHYTNWTHIYIC 17
          |: |||||||  ||:|
Db    124 DVHNTMHYTNWKFIYLC 140


RESULT 15
AAB98445
ID   AAB98445 standard; Protein; 353 AA.
XX
AC   AAB98445;
XX
DT   22-AUG-2001  (first entry)
XX
```

DE Human papillomavirus protein HPV33 E2.
XX
KW Human papillomavirus; human leukocyte antigen; HLA; immune response;
KW HPV; epitope; T cell; identification; vaccine; infection; genital wart;
KW neoplastic growth; antiviral.
XX
OS Human papillomavirus.
XX
PN WO200141799-A1.
XX
PD 14-JUN-2001.
XX
PF 11-DEC-2000; 2000WO-US33549.
XX
PR 10-DEC-1999; 99US-0172705.
PR 15-AUG-2000; 2000US-0641528.
XX
PA (EPIM-) EPIMMUNE INC.
XX
PI Sette A, Sidney J, Southwood S, Chesnut R, Celis E, Grey HM;
XX
DR WPI; 2001-381497/40.
XX
PT An isolated human papilloma virus (HPV) epitope, useful in vaccines for
PT treating HPV infections -
XX
PS Disclosure; Page 26; 756pp; English.
XX
CC The present invention describes an isolated prepared human papillomavirus
CC (HPV) epitope (I). (I) has antiviral activity, and can be used in
CC vaccine production. Peptides and corresponding nucleic acid compositions
CC from the present invention are useful for stimulating an immune response
CC to HPV by stimulating the production of CTL or HTL responses,
CC specifically in the treatment or prophylaxis of HPV infection, in persons
CC who have not manifested symptoms e.g. genital warts or neoplastic growth.
CC The peptides can also be used in a tetramer staining assay to assess
CC peripheral blood mononuclear cells for the presence of antigen-specific
CC CTLs following exposure to a pathogen or immunogen, and as reagents to
CC evaluate immune recall responses or evaluate the efficacy of a vaccine.
CC The vaccine compositions are useful for removing warts or treating HPV
CC infections. The epitopes for inclusion in an epitope-base vaccine may
CC be selected from conserved regions of viral or tumour-associated
CC antigens, which reduces the likelihood of escape mutants, also
CC immunosuppressive epitopes that may be present in whole antigens can be
CC avoided with the use of epitope-base vaccines. An additional advantage
CC is the ability to combine selected epitopes (CTL and HTL) and to modify
CC the composition of the epitopes achieving enhanced immunogenicity, the
CC major benefit of the vaccine is that is safe and efficacious. AAB98391
CC to AAB98477 represent polypeptide sequences used in the exemplification
CC of the present invention.
XX
SQ Sequence 353 AA;

Query Match 55.5%; Score 66; DB 22; Length 353;
Best Local Similarity 75.0%; Pred. No. 0.12;
Matches 12; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```
Qy       4 NTMHYTNWTHIYICEE 19
           ||| ||||  ||| ||
Db     127 NTMDYTNWGEIYIIEE 142
```

Search completed: January 14, 2003, 18:39:54
Job time : 37 secs